Fatigue Reliability of Ship Structures

Lei Yu

BEng, MSc

Submitted in fulfilment of the requirements for the

Degree of Doctor of Philosophy

Department of Naval Architecture & Ocean Engineering

Faculty of Engineering

University of Glasgow

October 2010

To my family

Abstract

Today we are sitting on a huge wealth of structural reliability theory but its application in ship design and construction is far behind. Researchers and practitioners face a daunting task of dove-tailing the theoretical achievements into the established processes in the industry. The research is aimed to create a computational framework to facilitate fatigue reliability of ship structures. Modeling, transformation and optimization, the three key elements underlying the success of computational mechanics are adopted as the basic methodology through the research. The whole work is presented in a way that is most suitable for software development.

The foundation of the framework is constituted of reliability methods at component level. Looking at the second-moment reliability theory from a minimum distance point of view the author derives a generic set of formulations that incorporate all major first and second order reliability methods (FORM, SORM). Practical ways to treat correlation and non-Gaussian variables are discussed in detail. Monte Carlo simulation (MCS) also accounts for significant part of the research with emphasis on variance reduction techniques in a proposed Markov chain kernel method. Existing response surface methods (RSM) are reviewed and improved with much weight given to sampling techniques and determination of the quadratic form. Time-variant problem is touched upon and methods to convert it to nested reliability problems are discussed.

In the upper layer of the framework common fatigue damage models are compared. Random process simulation and rain-flow counting are used to study effect of wide-banded non-Gaussian process. At the center of this layer is spectral fatigue analysis based on SN curve and first-principle stress and hydrodynamic analysis. Pseudo-excitation is introduced to get linear equivalent stress RAO in the non-linear ship-wave system. Finally response surface method is applied to this model to calculate probability of failure and design sensitivity in the case studies of a double hull oil tanker and a bulk carrier.

Summary

The opening pages put the topic of structural reliability in a wider context. Advancement of computational mechanics is discussed to draw useful lessons and analogies to shape philosophical and methodological foundation for the research.

Random number and random process generation techniques with FFT acceleration are introduced. Select-down of random number generator is made to facilitate a proposed Markov chain Monte Carlo method. It starts with an initial adaptive search to get the most probable failure point. This is followed by Markov chain pre-sampling to determine the kernel density. Kernel components are condensed with optimized global and local window width to reduce the squared error. Importance re-sampling is then carried out to calculate the final probability of failure.

To reiterate second-moment reliability theory a set of self-contained formulations is derived in terms of correlated variables of arbitrary distributions. Key algorithms are summarized in an easy to implement manner and numerical comparison is given between the main-stream FORM and SORM methods. Stepwise response surface is proposed to address cases where limit state function is only implicitly known. Terms included in the response surface function are determined by their statistical contributions. Overview of time-variant problem is given. It is shown that it can be converted to nested reliability problems at component or system level.

Fatigue damage mechanism is discussed in the forms of local strain, fracture mechanics and SN curve models. A new method to study the effect of wide-banded non-Gaussian stress response is proposed. It is based on rain-flow counting the full hysteresis cycles of random stress series generated according to distributions in Johnson family. Pseudo-excitation method is introduced into traditional spectral analysis to handle non-linearity due to external pressure in the splash zone and inertial loads. Frequency dependent pseudo wave heights are used to get linear equivalent RAOs as opposed to commonly used prescriptive wave heights. Stepwise response surface method is applied here to estimate fatigue reliability of typical structural details of a tanker and a bulk carrier model. Sensitivity analysis is performed to indicate the most efficient way to improve the design.

Contents

Acknowledgements

The progress of this research can be likened to the well known three stages of fatigue damage mechanism. The “initiation stage” was done during my stay at University of Glasgow. It all went well at first. But in the third year of my research I increasingly felt that the final answer to many problems I had encountered lied in the industry experience. I was lucky enough to get employed by Lloyd’s Register later that year and the “propagation stage” began. It was extremely painful in this stage to balance my work life, family life and the research work. But the constant support I got from people around me finally provided enough energy for me to reach the critical point for the final breakthrough.

I am enormously indebted to Professor Das, my supervisor. He is the one who opened the doors to the industry for me by providing loads of opportunities to communicate with people from outside the academia. He has showed great understanding towards my lack lustre progress after my work life began and never held back his best support to keep my research going.

I owe a lot to Dr Yunlong Zheng, a long time inspirer of mine, who has kindly provided me with steady encouragement whenever my spirit tends to wear out. I have tremendously benefited from his expertise in FE analysis and structural reliability over the years.

Professor Nigel Barltrop instrumentally shared his view on many occasions in the “initiation” years. His pursuit of simplicity in solving complicated engineering problems prompted me to think deeper and to attack problems from grass root level.

My special thanks are due to Mrs Thelma Will, the department research secretary. She truly cares about the wellbeing of every student and their family. She has kept me well informed of important matters related to my research. The new strength I found from her regular and warm compliments helped me weather many agonising moments.

I am extremely grateful for the development I have received from my current employer Lloyd’s Register. The experience I gained in my work has helped a lot with the finishing chapter of this thesis. I have been able to talk to many designers in shipyards and people from other classification societies. All this helped shape my view about practicality. I have also built up a

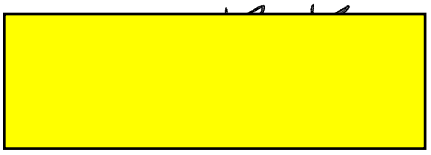
huge debt towards my colleagues for their knowledge sharing, in particular Dr Sai Wong and Dr Helena Polezhaeva.

The ultimate source of stamina has come from my family. My grandpa, who did not live to see the end of this research, offered a lot of encouragement. He taught me the importance of being a complete finisher from my childhood. I always think my grandma an embodiment of traditional Chinese virtues, especially self-esteem, self-reliance and perseverance. My parents are always supportive of my pursuits and are willing to provide the same old comfort zone whenever needed. I have been greatly enlightened by my father in law, a well known professor of Chinese philosophy. His influence helped me develop a universal mind towards life. Finally I would like to thank my wife and son wholeheartedly for putting up with my lousy role as a family man. To me their love and forbearance are the most certain things to count on in the world of uncertainty.

Lei Yu, London, May 2010

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Signature: 

Printed Name: Lei Yu

Chapter 1 Introduction

1.1 Objectives

For any emerging technology to be fully accepted by the industry collective intelligence must be developed first. Effective communication and knowledge transfer are the essential to help overcome the initial barriers. The whole process is highly iterative, and computational means must be provided to help grow the common experience in the new realm. This cannot be truer in ship building industry. The research is aimed to set up theoretical and computational framework for fatigue reliability of ship structures. This will be achieved though reiteration of traditional reliability theory from an angle in favor of real life implementation. Relevant numerical methods will be reviewed, compared and improved for better performance. The framework will be tested through predefined examples, and eventually, analysis of real ship structures in a double hull tanker and bulk carrier. It is certainly not possible to consider all the state-of-the-art techniques in the current research but the framework will be coherent and expandable so that new components or layers can be easily added to it.

1.2 Overview of the modern computational structural mechanics

The design of marine and offshore structures involves nearly every aspect of structural engineering. Fatigue reliability is merely a small part of it though by no means a trivial one. So it is sensible to set the whole work against the backdrop of the latest development of computational structural mechanics. In effect the prospect of any research in reliability being adopted by the industry greatly rests on how well it can leverage the wealth of existing framework of modern structural mechanics. System thinking plays a vital role here.

The last five decades have experienced a great revolution in structural design, which is triggered by the widespread application of numerical methods such as Finite Difference Method (FDM), Finite Element Method (FEM)[1,2] and Boundary Element Method (BEM)[3]. This has been accompanied by the surging performance and sliding cost of computation due to the rise of multi-processing technology. Though the theories of elastic, plastic and fracture mechanics have been established long since, not until the emergence of these numerical methods, in particular FEM and BEM, did they find profound use in engineering practice. These methods along with the ensuing variation principles and weighted residual theory constitute a brand new branch of mechanics, namely computational structural mechanics.

The versatility of computational structural mechanics set a vivid stage for engineers as well as researchers across different disciplines. In the mean time, to meet the ever increasing needs in engineering practice, the content of computational structural mechanics has been enriched dramatically. On the one hand techniques like pre-processing, solver, post-processing, and self-adaptation have helped static and dynamic analyses of deterministic structures reach so high a level that they have become indispensable in structural modeling and assessment [1]. On the other hand, because of the information and physical laws beyond our knowledge, existence of uncertainties in models we build to predict the behavior of real structural systems are irrevocable. Stochastic modeling methods are called upon where deterministic modeling is not sufficient or would give rise to misleading results [4]. In this case it is more reasonable and pragmatic to express the solution in a “weak form” based on probabilistic theory than to expect a traditional “precise solution”. In essence, this new thinking marks a considerable progress in structural mathematical modeling. Furthermore, the combination of computational structural mechanics with modern optimization and control theory is transforming the skyline of theoretic framework from an analytical one into a synthetic one [5,6].

As a matter of fact, modeling, optimization and control are closely related to each other. In general, optimization is the one element that penetrates through all the other components of modern computational structural mechanics. Not only does it underlie the variation principles and weighted residual method (this is quite straightforward if we take the energy functional and the residual of weight integration as the objective function), but it paves the path to

structural control. We might picture modern computational structural mechanics as a synthesis oriented framework in which modeling is the starting point and optimization is the core and sole means leading to control, the ultimate objective. Though these achievements are inspiring and the prospects of its application are inviting, the underlying framework is far from perfect now. Each part of it deserves further polishing. It should be noted that some achievements of modern mathematics have been introduced into the framework successfully and proven fruitful [7,8]. The following are those developments that have constructive bearing on naval architecture and ocean engineering.

1.2.1 Stochastic structural system

Perturbation finite element method is the dominant technique in this field and has made great achievements in static analysis, eigenvalue analysis, and structural stability analysis. This is greatly attributed to the establishment of random field theory [9]. However it behaves poorly in dynamic analysis and where the variation of basic random field is significant. As an alternative, stochastic simulation method is believed to be the most precise and versatile technique in nearly all kinds of probabilistic analyses. Unfortunately the prohibitive computational cost involved in complex structural system often degrades it into a method for verification in research.

Orthogonal polynomial expansion of basic random variable was a fad in the 90s. Typical of them are sequence orthogonal decomposition method and expanded order system method [4]. These methods prove flexible especially in compound stochastic vibration analysis. This reminds us of the great success achieved by Legendre transform in multivariate variation principle, which is nothing but an orthogonal transform. Orthogonization is one of the most efficient ways to decouple problems. It should be noticed that the stochastic modeling method using mean parameter identification has also been proposed [4].

1.2.2 Stochastic vibration

CQC (Complete Quadratic Combination) method is precise in theory. But the relevant computational cost is often intolerable. As an approximation, SRSS (Square Root of the Sum of Squares) method is usually recommended in which the cross-correlation items between the participant modes are generally neglected on the assumption that the damping is small and the participant eigen-frequencies are sparsely spaced. This assumption has proven vulnerable in practice.

One of the most noticeable progresses in this field is the application of pseudo excitation method (PEM) [10-18]. It takes into account the cross-correlation terms between the participant modes and between the excitations implicitly, and proves equivalent to CQC algorithm in nature while with much higher efficiency. In this method, the stationary multi-excitation problem as well as single-excitation problem can be calculated accurately by transforming the problem into the superimposition of a series of harmonic problems. In addition, the non-orthogonal damping matrix can be treated in a very convenient way without complex eigenvalue analysis. All these merits become even more distinct in non-stationary response analysis where the stochastic dynamic equations is transformed to a transient response equation with deterministic excitation to which the well-known numerical algorithms such as Newmark and Wilson-θ method and in particular some highly efficient time-step integration approaches are accessible [20-21]. PEM makes use of the finite element method to the full and is powerful in analysis of large-scale and complex structure systems. It is quite straightforward to develop new functions of stationary or non-stationary stochastic response on any existing programs with the basic function of harmonic or time variant analysis.

1.2.3 Optimization and control

These are the crown jewels of all the achievements in computational structural mechanics. Their synthesis and decision making nature greatly extended conventional structural mechanics theory. After nearly half a century's development structural optimization has won an important place in structural design. Because of its complexity any breakthrough in this field is often closely tied up with the methodology chosen. In general the methodology used

consists of modeling, transformation, optimization and inversion, of which modeling and transformation are the majority parts of the whole problem solving, accounting for 70%-80% of total work [6]. The methodology adopted here can be generalized to others fields, e.g. probabilistic analysis. This lends greatly to the current research.

The similitude relationship established between optimal control and computational structural mechanics indicates that control theory and computational mechanics are rooted in things of the same nature [5]. Indeed both theories can be derived from variation principle, for instance, the formulas of Kalman filter and H_∞ filter in system identification can be deducted easily by Lagrangian multiplier method. The importance of this finding hints that the techniques established in one of them can be applied to the other. Little wonder that subspace iteration method, conjugal subspace and conjugal symplectic subspace iteration method have found their use successfully in eigen-problem calculation of large control systems. Likewise multi-substructure techniques have proven very useful in nonlinear control systems.

1.3 State of art of structural reliability analysis

The theory of reliability analysis of structural systems consists of three parts: (1) Identification of dominant failure mode; (2) Calculation of probability of failure and sensitivity with respect to the design parameters; (3) Determination of the upper and lower bounds of the overall structural system according to the correlation between the dominant failure modes and their probability of failure. Of these, part (2) and (3) are at the centre of the whole reliability theory.

1.3.1 Dominant failure mode

The identification of dominant failure mode is usually performed by mechanical methods [22] or mathematical programming. Though the proposed methods all show some kind of effectiveness, the relevant computation is still taxing for complex structural systems. In fact the true problem in this case is that no closed form limit states can be obtained because numerical methods are used to calculate the response. This really raises a hurdle in way of its application in engineering practice. But otherwise it happens to provide another vivid example that any development in structural analysis is greatly dependent on the advances of

computational mechanics. As a compromise, people are often content to study structural components or its local behavior.

1.3.2 Second moment theory

Calculation of probability of failure is the most mature part in reliability theory. Thanks to the foundation work of Cornell (1969), Hasofer and Lind (1974) and Shinozhka (1983) [23-25] et al., first order and second order second moment theories (FOSM and SOSM) are well established and have found ever-increasing use in different engineering fields [26-29]. It should be noted that, in this theory the integration of joint probability function of design variables is circumvented by transforming it into a least distance problem in standard normal space. Orthogonal transform is again used to uncouple the correlated design variables. This shows that the essence of reliability analysis is nothing but optimization. Hence, many optimization theory and algorithms in mathematics are readily applicable in reliability analysis, which dramatically enriches the computational methods in this realm. And we cannot but realize again the marvelous similarity between the development of reliability analysis and that of computational structural mechanics. The enlightenment prompts us to choose appropriate methodology spontaneously in our research.

1.3.3 Reliability of structural systems

After the design points are obtained through first order methods series and parallel structural system reliability can be reduced to evaluation of multi-normal integral. The integral itself is an intersection problem and can only be approximated when the dimension is higher than 2. Because of the computational cost in problems of high dimension, the failure probability of structural systems often comes in a “weak form”. Instead of calculating the probability of failure itself bounding technique can be used. Two useful formulations are wide bound method [30] and narrow bound method [31]. They are first order and second order approximation respectively. However with the increase of failure modes and their correlation, the bounds will become too loose. In this case, formulation of higher order approximation can be developed on

the one hand. On the other hand, different point evaluation technique can be used [32,33]. Some modified bound methods and point evaluation methods can be found in [29].

1.3.4 Monte Carlo simulation

Monte Carlo simulation (MSC) also plays a very important role in different levels of reliability analysis. The accuracy it provides is only dependent on the number of sample points while not affected by the distribution type and the number of basic variables. What's more it can be used in those cases where the limit state function is not known implicitly and it is the only approach to the highly nonlinear problems. Despite the introduction of variance reduction techniques such as importance sampling method, the computation cost in large scale and complex structural system analysis is still formidable to outweigh its other advantages. For all these reasons, MSC has long been looked on as Cinderella in the field of probabilistic analysis, who is waiting for the Princess armed with cheaper and more powerful computing techniques.

1.3.5 Response surface method

Response surface method (RSM) is one of the most sparking developments in structural reliability analysis. It is highly suitable for the case where the limit state function has no known closed form and must be evaluated point-wisely by numerical methods such as FEM. In a sense, RSM is a system identification procedure, in which a transfer function relating the input parameters (loading and system conditions) to the output parameters (response in terms of displacements, stress, etc.) will be found in a suitable way. The observations required for the identification are usually taken from systematic numerical experiments with the full mechanical model, and the transfer function obtained approximately is termed response surface function (RSF). The ideology of RSM can be traced back to the fifties in experiment field, but only recently was it introduced into the field of structural reliability analysis. It brings a touch of freshness to the reliability theory by rendering possible a perfect combination of the deterministic structural analysis software and the basic reliability method aforementioned. On the other hand, even for those problems accessible to other types of approximations, the RSM is shown to be superior in both accuracy and efficiency. The key

problems in RSM are the experiment design and the identification of unknown parameters in RSF. They decide completely the accuracy and efficiency of the whole algorithm. The work of Bucher (1990) and Rajashekhar (1993) set the tone of RSM [34,35]. Typical of the following advanced algorithms can be found in [36]-[38]. Particularly, a very thorough adaptive iteration scheme is proposed in [38] by weighted step-wise regression, which will be applied in the current research.

1.4 Application of reliability fatigue analysis to marine structures

Fatigue is one of the most common damages in marine structures. Since the damage is caused by the accumulated effect of continuous cyclic loads rather than the once-in-life ultimate loads, it leads to structure failure at much lower stress level than the allowable stress by other strength criteria. Besides, it tends to trigger other kinds of structure failure at low stress level, such as brittle fractures and buckling.

1.4.1 Cumulative fatigue damage model

On the enlightenment of damage accumulation concept introduced by Palmgren about 70 years ago and 'linear damage rule' by Miner about 50 years ago, a lot of fatigue damage models has been published, which fall into the following categories: [39]

- Linear damage evolution and liner summation
- Nonlinear damage curve and two-stage linearization approaches
- Life curve modifications to account for load interactions
- Approaches based on crack growth concept
- Models based on continuous damage mechanics (CDM)
- Energy-based methods
- Fatigue reliability models

The tendency of the fatigue damage modeling is from empirical, phenomenological, and deterministic models to analytical and stochastic models. This is the inevitable result of the

significant progress in computational mechanics, damage mechanics, and experimental techniques.

Despite the above achievements cumulative fatigue is far from resolved. No model proves to be a panacea, rather they are all restricted to some particular cases. It is insurmountable to propose a method that can take into account of all the factors in fatigue damage. In addition though the proposed nonlinear models look pretty conceptually, there is no significant evidence that they are superior to the linear models. As a consequence, the models most favored by the practitioners are still those semi-empirical models, like S-N curves and Paris law, based on linear damage accumulation theory. However, it is of vital importance to tailor any models according to practical problem to prevent misuse [39]. Typical of fatigue analysis approaches for ship structures are those proposed by Petinov [41,42], in which S-N curve in experiments is modified from different respect to account for the mismatch between test specimen and practical structural detail. In the hybrid SN-FM model proposed by Xu[43,44], effects of weld and loading modes are studied and the concept of S-N curve is extended to structural details with crack.

1.4.2 Contributors to fatigue damage

The variety of fatigue damage models stems from the complexity of fatigue phenomenon in engineering practice. The following is a list of the common factors that can affect fatigue strength:

I. Workmanship
- Fabrication of structural details
- Weld geometry and defects
- Residual stress
- Heat-affected zone
- Fusion zone material

II. Loading Condition

- Multi-axial random load
- Load history
- Load frequency
- Mean load
- Overload
- Load shedding

III. Crack Propagation Mechanism

- Crack closure
- Scale effect in application of experiment results
- Crack growth threshold
- Plasticity in high-speed propagation

IV. Environment

- Corrosion
- Other factors leading to initial crack

1.4.3 Spectral fatigue analysis

Spectral analysis is now commonplace when calculating stress responses of a linear structure-wave system. Since any random wave can be decoupled into the superimposition of a series of regular waves, structural responses such as stress only need to be evaluated once for unit wave amplitude. The moments of the stress spectrum can be used to determine the parameters of a predefined PDF for stress. Then different damage accumulation models can be applied to predict fatigue life [45,46]. Alternatively the stress spectrum obtained can be used to generate time series of stress process. Along with MSC simulation and rain flow method, it is possible to consider the effect of bandwidth on fatigue damage, to correct the otherwise conservative estimation by narrow-banded assumption [47].

One of the biggest limitations of spectral analysis is that it is not applicable in nonlinear system. In this case, transient analysis in time domain might be considered. Static analysis or

any of the numerical integration routines mentioned in 1.1 can be used to calculate the time series of response. In reality because the computation cost in transient analysis is much higher than spectral analysis an equivalent linear transfer function is often preferred to obtain the response spectrum. This can be performed by different equivalent linearization methods. Problems that arise here are studied systematically by Bishop (1996) [48]. Another compromise used is nonlinear transfer function approach, which only involves limited time-domain analysis [49]. As a matter of fact pseudo-simulation method is a very promising choice in this field because of its numerical efficiency.

1.4.4 Fatigue reliability

Fatigue strength is sensitive to the uncertainties of the structural system. In fact the dimension (i.e. the size of local plastic zone) governing the fatigue behavior is so small that it is the totality of randomness that dominates the macro phenomena. All this makes stochastic fatigue modeling and fatigue reliability analysis more appealing.

Just as in other fields of reliability analysis, most of the random models in fatigue reliability analysis are obtained by simply randomizing the relevant parameters of their counterparts in the deterministic realm. Sensitivity analysis often plays an important role in this process to identify the importance of each parameter. A good summary of methods available can be found in [27]. A plethora of other models have also been developed with specific reliability techniques [50-52].

Since the structural reliability theory has been well established, now the main problems in reliability analysis rest with the incomplete knowledge of distributions and statistics of the design variables, especially those related to loads and responses. If these properties are poorly assumed the corresponding reliability analysis would not contribute to effective decision making. No reliability analysis should be done for its own sake without observing important prerequisites. It is the very time to establish a rational database by the joint efforts of the research, design, construction, and regulatory bodies. Fortunately, this has caused enough attention of many researchers. One of the themes in their work is to fit the parameters of

assumed distribution or to formulate particular distributions by delicate probabilistic analysis and stochastic simulation [53,54].

1.4.5 Fatigue damage control

The accuracy of reliability analysis of any level relies greatly on how much information has been obtained about a particular problem and the way the information is organized. In fatigue analysis, it is far from enough to use only those pieces of information obtained in design stage. To add more confidence to the prediction, the model established in the design stage must be updated according to the actual state of structure in service. This involves development of a procedure of system identification along with a rational inspection program, which is the necessary way leading to fatigue control. In addition, the planning of inspection, modification, and repair are also an important part of fatigue control. Since these are inevitable in practice, an appropriate overall strategy in early stage could save considerable time and expenses.

Bayesian analysis is usually used in reliability model identification. According to the different part to be identified, it falls into three categories, namely, event updating, POD (Probability of Detection) updating, and Bayesian estimation. In event updating, the failure probability itself is updated. In POD updating, the multi-variant probability distribution is identified. While in Bayesian estimation, the statistics of design variable are identified. The latest application of these can be found in [55]. Some problems in the algorithm of updating are discussed thoroughly in [56].

1.5 About the thesis

This research is aimed at developing systematic methods in fatigue reliability analysis of ship structures based on the latest achievements in relevant fields. Techniques that the author thinks most relevant to fatigue reliability for ship structures are carefully chosen and discussed. The basic theory is covered in such a way that the mathematical deduction is done where necessary to put the traditional theory in a context favorable to computer programming.

While oil tanker and bulk carrier will be given particular consideration, the proposed methods are generic enough to be applied to other marine structures without overhaul. For the sake of feasibility and uniformity, the methodology used will keep abreast of the trend of the development of modern computational mechanics.

Chapter 2 introduces the concept of random number generation and its use in simulating random processes. The performance of a set of basic generators is compared in detail. Down selection is made for the following reliability study in this thesis. Random process generation based on FFT is discussed.

Chapter 3 establishes the second moment reliability formulations in the most generic form. First and second order reliability methods (FORM and SORM) are re-interpreted in an optimization context step by step. Relations between the incumbent FORM and SORM methods are fully revealed. The formulations are self contained with key algorithms summarized to assist practical use in solving engineering problems.

Chapter 4 presents principles of Monte Carlo simulation in reliability analysis. A new adaptive importance sampling method based on kernel density estimate is proposed. In construction of the kernel density, a weighting coefficient is assigned to each component. These coefficients along with the position of each kernel component are refined in the following adaptive iteration, which is initialised by a Markov chain. Uniformity entropy in terms of the weighting coefficients is calculated to monitor the goodness of fit. To reduce the computational time in iteration an analytical approximation is used as the optimum window width. The fitted kernel mixture will be clustered to reduce number of components with the window width optimised either through cross validation or analytical approximation.

Chapter 5 discusses the fundamentals of response surface method (RSM). A stepwise response surface approach is then proposed. By means of stepwise regression the square and cross terms in the quadratic polynomial can be picked up automatically according to their actual contributions in variance analysis. The down selection can be further controlled by applying weighting factors to the statistical value of each term's contribution and changing the

thresholds for acceptance and rejection. Goodness of fit is measured by well proven criteria in the traditional statistical tests. The whole algorithm starts with a linear response surface. As the adaptive iteration proceeds, the bar on quadratic terms is lifted gradually to allow them to enter the model. Since the sample points in one step of iteration are recycled in the subsequent ones, a simple experimental design is enough to fit a robust response surface. A double bottom hull system is analyzed with randomized Young's modulus, load distribution, and geometric properties. Results are compared with direct FORM, SORM and MCS.

Chapter 6 looks at the fatigue damage models based on SN curve, local strain and fracture mechanics. Their applicability in different crack stage is examined. Influential factors in fatigue damage are discussed. This is followed by derivation of a generic reliability model. A new attempt is made in this chapter to quantify the effect of bandwidth and non-normality in fatigue damage analysis. For the lack of actual stress history, a series of non-Gaussian and homogeneous random processes are generated with fast Fourier transform (FFT) acceleration. A correction factor is defined on the basis of rain-flow counting. It is revealed that the fatigue damage evaluated through traditional SN approach may be either conservative or rather non-conservative. The upper and lower bounds of the correction factor are studied with respect to kurtosis and skewness of the generated random process and the slope of SN curve.

Chapter 7 gives a detailed literature review on the first passage problem. This is followed by derivations to show how nested reliability models can be used to calculate the failure probability at given time T. Its potential application to fatigue reliability is touched upon using a time variant model based on fracture mechanics.

Chapter 8 attempts to introduce structural uncertainties in established fatigue design assessment process such as ShipRight FDA3 procedure and the supporting software by Lloyd's Register. Key design parameters are randomized in a spectral fatigue model, where pseudo-excitation method is used to consider the nonlinear effect of inertial loads and external wave pressure in the splash zone. Step-wise response surface method is used in tandem with t--by-t fine mesh FE analysis to obtain the probability of failure. The calculation is demonstrated in two case studies for an oil tanker model and a bulk carrier model respectively.

Chapter 9 concludes the current work with a prospect for future research. A list of publications out of this research is given at the end.

References

1. Zienkiewicz, O.C, Taylor, P.L. (1989). *Finite Element Method*, London: McGRAW-HILL.

2. Cook, R. D., Malkus, D. S., & Plesha, M. E. (1989) *Concepts and Applications of Finite Element Analysis*, Chichester: Wiley.

3. Brebbia, C.A. (1978). *Boundary Element Method for Engineers*, London: Pentech Press.

4. Li, J. (1996). *Stochastic Structural System: Analysis and Modeling*, Beijing: Science Press.

5. Zhong, W. X. (1993). *Computational Structural Mechanics and Optimal Control*, Dalian: Dalian University of Technology Press.

6. Sui, Y. K. (1996). *Modeling, Transformation and Optimisation*: *New Developments of Structural Synthesis Method*, Dalian: Dalian University of Technology Press.

7. Arnold, V. I. (1978). *Mathematical Methods of Classical Mechanics*, New York: Springer-Verlag.

8. Zhong, W. X., & Williams, F. W. (1992). Wave problems for repetitive structures and symplectic mathematics. *Proceedings of the Institution of Mechanical Engineers-Journal of Mechanical Engineering Science*, 206(6), 371-379.

9. Vanmarke, E. (1983). *Random Fields: Analysis and Synthesis*, Cambridge: MIT Press.

10. Lin, J. H. (1985). A deterministic algorithm for stochastic seismic responses. *Chinese Journal of Earthquake Engineering and Engineering Vibration*, 5, 89-94.

11. Lin, J. H. (1992). A fast CQC algorithm for random seismic responses. *Computers & Structures*, 44(3), 683-687.

12. Lin, J. H. (1993). An accurate and highly efficient algorithm of non-stationary stochastic seismic responses. *Chinese Journal of Earthquake Engineering and Engineering Vibration*, 13, 89-94.

13. Lin, J. H., Williams, F. W., & Zhang, W. S. (1993) A new approach to multi-phase excitation stochastic seismic response. Microcomputers in Civil Engineering, 8(4), 283-290.

14. Lin, J. H., Zhang, W. S., & Williams, F. W. (1994). Psedo-excitation algorithm for nonstationary random seismic response. *Engineering Structures*, 16(4), 270-276, 1994.

15. Lin, J. H., Zhang, W. S. & Li J. J. (1994) Structural responses to arbitrarily coherent stationary random excitations. *Computer & Structures*, 50(5)**,** 626-623.

16. Lin, J. H., Zhi, H., & Guo, X. L. (1998) Inverse pseudo excitation method for loading identification of stationary random vibration (1). *Chinese Journal of Computational Mechanics*, 15(2), 127-136.

17. Zhong, W. X. (1996). A series of highly efficient approaches to structural stochastic response. *Proceedings of Natural Science-Transaction of State Key Laboratory*, China, 6(4), 394-401.

18. Lin, J. H., & Zhong, W. X. (1998). Some notes on FEM and structural random response analysis. *Chinese Journal of Computational Mechanics*, 15(2), 217-223.

19. Zhong, W. X. (1994). On precise time-integration method for structural dynamics. *Journal of Dalian University of Technology*, 34(2), 131-136.

20. Zhong, W. X., & Williams,_F. W. (1994) Precise time step integration method. *Proceedings of the Institution of Mechanical Engineers-Journal of Mechanical Engineering Science*, 208(6), 427-430.

21. Lin, J. H. (1995). Compound precise time-integration approach to for structural non-stationary stochastic response. *Chinese Journal of Vibration Engineering*, 8(2), 127-134.

22. Thoft-Christensen, P., & Murotsu, Y. (1986). *Application of Structure Systems Reliability Theory*. Tokyo: Springger-Verlag.

23. Cornell, C. A. (1969) A Probability-Based Structural Code. *Journal of the American Concrete Institute*, 66(12), 974-985, 1969.

24. Hasofer, A. M., & Lind, N. C. (1974) Exact and invariant second-moment format. *Journal of the Engineering Mechanics Division*, ASCE, 100(1), 111-121.

25. Shinozuka, M. (1983). Basic Analysis of Structural Safety. *Journal of Structural Engineering*, 109(3), 721-740.

26. Ang, A. H. -S., & Tang, W. H. (1984). *Probability Concept in Engineering Planning and Design*, New York: Wiley.

27. Madsen, H.O., Krenk, S., & Lind, N.C. (1986). *Methods of Structural Safety*, New Jersey: Prentice-Hall.

28. Ditlevsen ,O., & Madsen, H. O. (1996). *Structural Reliability Methods*, Chichester: John Wiley & Sons.

29. Zhao, G. F. (1996) *Reliability theory and its applications for engineering structures*. Dalian: Dalian University of Technology Press.

30. Cornell, C. A. (1967). Bounds on the reliability of structural systems. *Journal of the Structural Division*, ASCE, 93(ST1), 171-200.

31. Ditlevsen, O. (1979) Narrow reliability bounds for structural systems. *Journal of Structural Mechanics*, 7(4), 453-472.

32. Ang, H-S., & Ma, H. F. (1981) On the reliability of structural systems. *Proceedings of International Conference on Structural Safety and Reliability*, Trondheim.

33. Song, B. F. (1992). A numerical method in affine space and a method with high accuracy for computing structural system reliability. *Computers & Structures*, 42(2), 255-262.

34. Bucher, C. G. & Bourgund, U. (1990) A Fast and Efficient Response Surface Approach for Structural Reliability Problems. *Structural Safety*, **7**(1), 57-66.

35. Rajashekhar, M. R., & Ellingwood, B. R. (1993). A New Look at the Response Surface Approach for Reliability Analysis. *Structural Safety*, 12(3), 205-220.

36. Kim, S. H., & Na, S. W. (1997) Response surface method using vector projected sampling points. *Structural Safety*, 19(1), 3-19.

37. Zheng, Y. L., & Das, P. K. (2000) Improved response surface method and its application to stiffened plate reliability analysis. *Engineering Structures*, 22(5), 544-551.

38. Yu, L., Das, P. K. & Zheng, Y. L. (2002). Stepwise response surface method and its application to reliability analysis of ship structures. *Journal of Offshore Mechanics and Arctic Engineering*, 124(4), 226-210.

39. Fatemi, A. & Yang L. (1998). Cumulative fatigue damage and life prediction theories: a survey of the state of the art for homogeneous materials. *International Journal of Fatigue*, 20(1), 9-34.

40. Paris, P. C. (1998). Fracture mechanics and fatigue: a historical perspective. *Fatigue & Fracture of Engineering Materials & Structures*, 21(5), 535-540.

41. Petinov, S. V. (1990). *Basic Engineering Analysis of Fatigue of Ship Structures*, St. Petersburg: Sudostroenie Publishing.

42. Petinov, S. V., & Thayamballi, A. K. (1998). The application of S-N curves considering mismatch of stress concentration between test specimen and structure. *Journal of Ship Research*, 42(1), 68-78.

43. Xu, T., & Bea, R. G. (1997). Fatigue of ship critical structural details. *Journal of Offshore Mechanics and Arctic Engineering*, 119(2), 96-107.

44. Xu, T. (1997). Fatigue of ship structural details-technical development and problems. *Journal of ship research*, 41(4), 318-331.

45. Chen, Y. N., & Mavrakis, S. A. (1988). Closed-form spectral fatigue analysis for compliant offshore structures. *Journal of Ship Research*, 32(4), 297-304.

46. Det Norske Veritas. (1998). *Fatigue assessment of ship structures*, Oslo: DNV.

47. Wirsching, P. H., ASCE, A. M., & Light, M. C. (1980). Fatigue under wide band random stresses. *Journal of structural division*, ASCE, 106(ST7), 1593-1607.

48. Bishop, N. W. M., Schofield, P., Kirkwood, M. G., Turner, T. (1996) Spectral fatigue analysis of shallow water jacket platforms. *Journal of Offshore Mechanics and Arctic Engineering*, 118(3), 190-197.

49. Jha, A. K., & Winterstein, S. R. (2000). Stochastic fatigue damage accumulation due to nonlinear ship loads. *Journal of Offshore Mechanics and Arctic Engineering*, 122(4), 253-259.

50. Lanning, D., & Shen, M.-H. H. (1998). Reliability of welded structures containing fatigue cracks. *Journal of Offshore Mechanics and Arctic Engineering*, 118(4), 300-306.

51. Garbatov, Y., & Soares, C. G. (1998). Fatigue reliability of maintained welded joints in the side shell of tankers. *Journal of Offshore Mechanics and Arctic Engineering*, 120(1), 2-9.

52. Hu, Y. R., Chen, B. Z., & Ye, N. Q. (1998) Fatigue reliability analysis of redundant structural systems by using Monte Carlo simulation. *Offshore Mechanics and Arctic Engineering*.

53. Huijsmans, R. H. M., & Adeegeest, L. J. M. (1998). Fatigue assessment in FPSO mooring design using moment based Hermite approximation. *Offshore Mechanics and Arctic Engineering*.

54. Lanning, D. B. Jr., & Shen, M.-H. H. (1998) Reliability of welded structures containing cracks in heat affected zones. *Offshore Mechanics and Arctic Engineering*.

55. Moan, T., & Song, R. (1998). Implication if inspection updating on system fatigue reliability of offshore structures. *Offshore Mechanics and Arctic Engineering*.

56. Sindel, R., & Rackwitz, R. (1998). Problems and solution strategies in reliability updating. *Journal of Offshore Mechanics and Arctic Engineering*, 120(2), 109-114.

Chapter 2 Random Process Generation

2.1 Introduction

Probabilistic analysis of structural systems relies heavily on availability of statistical information. In most of the cases second order information is sufficient, while in certain types of fatigue analysis information of higher order is needed. Because of the complexity in this field, information available through analytical tools is very limited. As a result, Monte Carlo simulation (MCS) is often resorted to not least because of its high accuracy and ease of implementation. The only drawback of MCS is always associated with its high computational cost. But this is being removed by cheaper and fast advancing computing technology.

In the first part of this chapter, a few basic yet efficient algorithms in random number generation and test are introduced. This is followed by a detailed discussion on the random process generation technique based on spectrum representation and FFT. Numerical examples are presented to conclude our discussion.

2.2 Random number generation

Random number generation is the centrepiece of Monte Carlo simulation (MCS). In theory, if a sequence of random number uniformly distributed in the interval [0,1] is obtained, it can be transformed to another sequence with any other arbitrary distribution. So random generator for uniform distribution only is discussed here.

Nearly all the random generators are based on a recursive function. By far the most commonly used generator is linear congruential generator (LCG). The underlying recursive function is

$$\begin{aligned} i_{n+1} &= (ai_n + b) \bmod m \\ x_{i+1} &= i_{n+1} / m \end{aligned} \tag{2.1}$$

A seed i_0 is needed to start the procedure. The choice of multiplier a, increment b, and modulus m is dependent on the word length of computer w and it affects the quality of generated random number considerably. The principles to observe in performing this can be found in literature by Knuth (1981) and Rubinstein (1981). In addition, the generator must pass a series of statistical tests. One famous branch of LCG is prime modulus multiplicative generator, in which m is a prime number and $b = 0$. It proves not inferior to the other kinds of LCG, if not better.

The implementation of Eq. (1) often involves the multiplication of two 32-bit integers modulo a third 32-integer. To prevent the intermediate result from overflowing on a 32-bit computer, and to make the generator portable, the right hand side of Eq. (1) can be calculated by Schrage's trick (1979), which, for a multiplicative congruential generator, is

$$ai_n \bmod m = \begin{cases} a(i_n \bmod q) - \lfloor i_n/q \rfloor r & \text{if not negative} \\ a(i_n \bmod q) - \lfloor i_n/q \rfloor r + m & \text{otherwise} \end{cases} \tag{2.2}$$

where

$$q = \lfloor m/a \rfloor, \text{and } r = m \bmod a \tag{2.3}$$

so that

$$m = aq + r \tag{2.4}$$

Obviously a mathematical generator like LCG can only produce pseudo-random numbers rather than true ones. There are correlations implicitly imbedded between the numbers generated. To alleviate its side effect, the order of the random number can be rearranged by a standard shuffling scheme due to Bays and Durham (1976). The procedure is as follows:

- Create an array A of N integers between 0 and m-1 by a LCG with period m. N is typically around 100 and the exact number is of no importance. Then generate an additional integer y.

- Calculate the pointer of shuffling array $k = \lfloor yN/m \rfloor$, which is an integer between 0 and $N-1$.
- Return A(k) as the new random number of shuffling algorithm, and set y = A(k).
- Generate a new number by LCG and store it in A(k).

Combined together the above techniques can give generators more than sufficient for structural reliability analysis. For random number of higher quality, generators based on other recursive functions, such as shift register generator and lagged Fibonacci generator, can be used (Newman and Barkema 1999). But they are more complicated and over-qualified for our problem.

2.3 Statistical test of random number

No generator can be used without passing necessary statistical test. The general idea of statistical testing is to construct a statistic under certain hypothesis H_0, and calculate its value F using the sample obtained. Given significance level α, if F is greater than the critical value F_α then H_0 is denied. The following are some necessary tests for uniformly distributed random number.

Parameter test is used to check if the parameters of empirical distribution and those of theory distribution are different significantly. The statistics of the mean value, mean square and variance of uniformly distributed samples are

$$\begin{aligned} \bar{x} &= \frac{1}{N}\sum_{i=1}^{N} x_i \\ \bar{x^2} &= \frac{1}{N}\sum_{i=1}^{N} x_i^2 \\ s^2 &= \frac{1}{N}\sum_{i=1}^{N}(x_i - \frac{1}{2})^2 = \bar{x^2} - \bar{x} + \frac{1}{4} \end{aligned} \tag{2.5}$$

with probabilistic parameters

$$
\begin{aligned}
E[\bar{x}] &= \frac{1}{2}, D[\bar{x}] = \frac{1}{12N} \\
E\left[\bar{x^2}\right] &= \frac{1}{3}, D\left[\bar{x^2}\right] = \frac{4}{45N} \\
E[s^2] &= \frac{1}{12}, D[s^2] = \frac{1}{180N}
\end{aligned}
\tag{2.6}
$$

where E[·] is the operator of mathematical expectation and D[·] is the operator of variance. According to the law of large number we have

$$
\begin{aligned}
u_1 &= \frac{\bar{x} - E(\bar{x})}{\sqrt{D(\bar{x})}} = \sqrt{12N}(\bar{x} - \frac{1}{2}) \\
u_2 &= \frac{\bar{x^2} - E(\bar{x^2})}{\sqrt{D(\bar{x^2})}} = \frac{1}{2}\sqrt{45N}(\bar{x^2} - \frac{1}{3}) \\
u_3 &= \frac{s^2 - E(s^2)}{\sqrt{D(s^2)}}\sqrt{180N}(s^2 - \frac{1}{12})
\end{aligned}
\tag{2.7}
$$

which are asymptotic standard normal variables.

Uniformity test is used to check if the difference between empirical frequency and theoretical frequency are significant. If we divide interval [0,1] into k equal subintervals, then the theoretical number in each subinterval will be $m_j = N/k$. Assume n_j random numbers fall into the j_{th} subinterval, central limit theory gives the following statistics

$$
\chi^2 = \sum_{j=1}^{k} \frac{(n_j - m_j)^2}{m_j} = \frac{k}{N}\sum_{j=1}^{k}(n_j - \frac{N}{K})^2 \tag{2.8}
$$

which has $\chi^2(k\text{-}1)$ distribution asymptotically.

Independence test is used to check if the correlation between the random numbers is significant. One way to achieve this is to calculate the coefficient of auto-correlation with respect to lag j

$$
\hat{\rho}_j = \left[\frac{1}{N-j}\sum_{i=1}^{N-j} x_i x_{i+j} - (\bar{x})^2\right] \Big/ s^2 \tag{2.9}
$$

If N is large enough (say N- j > 50), statistics

$$u = \hat{\rho}_j \sqrt{N - j} \tag{2.10}$$

has standard normal distribution asymptotically, under the hypothesis $\rho = 0$. Usually, for j =1(1)20, u should not exceed the fractile under given significant level $F_{1-\alpha/2}$ more than twice.

2.4 Descriptions of random series

A realisation of random time series $x(t)$ can be regarded as an indexed family of random variables denoting the collective outcome of all the experiments comprising the random process. Hence all descriptions of random variable can be applied to random process as well (Papoulis 1965, Vanmarcke 1983).

The mean value $m(t)$ of $x(t)$ are named ensemble mean. The autocovariance function is defined as

$$C(t_1, t_2) = E[x(t_1)x(t_2)] - m(t_1)m(t_2) \tag{2.11}$$

The corresponding coefficient of correlation is

$$\rho(t_1, t_2) = \frac{C(t_1, t_2)}{\sqrt{C(t_1, t_1)C(t_2, t_2)}} \tag{2.12}$$

The covariance function of zero mean random fields is termed autocorrelation function and is denoted by $R(t_1,t_2)$ hereafter. Equation (11) and (12) are the second order description of random process. Both C and ρ are self-joint and positive definite calculus. They comprise all the information of a Gaussian process. For non-Gaussian process they are still used as the criterion to examine the convergence of simulated time series.

A random process is stationary (homogeneous) if its probabilistic properties are shift-invariant. So we have

$$\begin{aligned} m(t) &= m \\ C(t_1,t_2) &= C(t_1-t_2) = C(\tau) \\ \rho(t_1,t_2) &= \rho(t_1-t_2) = \rho(\tau) \\ R(t_1,t_2) &= R(t_1-t_2) = R(\tau) \end{aligned} \tag{2.13}$$

A random process is ergodic if its probabilistic properties can be represented fully by a single realisation. A random process is ergodic in the mean if

$$E[x(t)] = m(t) = \langle x(t) \rangle \tag{2.14}$$

where operator E[·] is the ensemble expectation across all realisations at given time t, operator $\langle \cdot \rangle$ is the temporal average of a single realisation, i.e.

$$\langle x(t) \rangle = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)dt \tag{2.15}$$

Likewise, the ergodicity in the correlation is defined by

$$E[x(t)x(t+\tau)] = \langle x(t)x(t+\tau) \rangle \tag{2.16}$$

It can be readily seen that if a random process is ergodic, it must be stationary as well. However this is not true the other way around.

If there is no harmonic components in $x(t)$, the autocorrelation function $R(\tau)$ is L^2 integrable (it decays when τ increases), which yields the following Fourier transform pair

$$\begin{cases} S(\omega) = \dfrac{1}{2\pi} \displaystyle\int_{-\infty}^{\infty} R(\tau) e^{-i\omega\tau} d\tau \\ R(\tau) = \displaystyle\int_{-\infty}^{\infty} S(\omega) e^{i\omega\tau} d\omega \end{cases} \tag{2.17}$$

where $S(\omega)$ is the spectra density function (s.d.f.) of random field. This is the famous Wiener-Khinchine relation. It is of vital importance in random field theory and application.

2.5 Generation of Gaussian random process

In general, a realisation of random field can be simulated by means of:

1) Spectral representation method

2) ARMA (auto-regressive moving average) modelling

3) covariance matrix decomposition method

4) Scale refinement method

5) Shot noise, random Fourier series, and noise shower processes

Each method possesses some advantages and disadvantages. Which to choose depends on particular problem in hand. A wide prospective of the above methods can be found in the paper by Spanos and Zeldin (1998).

In structural stochastic analysis, spectral representation distinguishes itself from the others by its high efficiency, and is widely adopted. According to Shinozuka (1991), a stationary and ergodic random time series $x(t)$ with zero mean, and two-sided power spectral density $S(\omega)$ can be simulated by the following expansion

$$x(t) = \sqrt{2}\sum_{0}^{N-1} A_n \cos(\omega_n t + \Phi_n) \quad N \to \infty \tag{2.18}$$

where Φ_n are independent random phase angles distributed uniformly over $[0, 2\pi]$ and the amplitudes can be expressed by

$$\begin{gathered} A_n = (2S(\omega_n)\Delta\omega)^{1/2}, \; n = 0,1,2,\ldots,N\text{-}1 \\ \omega_n = n\Delta\omega, n = 0,1,2,\ldots,N\text{-}1 \\ \Delta\omega = \omega_u / N \end{gathered} \tag{2.19}$$

of which

$$A_0 = 0 \text{ or } S(\omega_0 = 0) = 0 \tag{2.20}$$

In Eq. (19) ω_u is the upper cut-off frequency beyond which the value of $S(\omega)$ is negligible. According to Conte and Boor (1980), the error of ensemble and temporal autocorrelation function of simulated process can be evaluated by

$$\begin{aligned} Error &= \int_0^{\omega_u} 2S(\omega)\cos(\omega\tau)d\omega - \sum_{n=0}^{N-1} 2S(\omega_n)\cos(\omega_n\tau)\Delta\omega \\ &= \frac{\omega_u^2}{2N}[\frac{d}{d\omega}(S(\omega)\cos\omega\tau)]\Big|_{\omega^*} \end{aligned} \tag{2.21}$$

where ω^* is a intermediate value between 0 and ω_u. This indicates that the error between simulated ensemble autocorrelation and the target autocorrelation is proportional to $1/N$. Further, if we define ω_n as

$$\omega_n = n\Delta\omega + \Delta\omega/2, \quad n = 0,1,2,\ldots,N-1 \tag{2.22}$$

Eq. (21) will be rewritten as

$$Error = \frac{\omega_n^3}{24N^2}[\frac{d^2}{d\omega^2}(S(\omega)\cos\omega\tau)]\Big|_{\omega^*} \tag{2.23}$$

Now the error becomes proportional to $1/N^2$. However the use of Eq. (22) will prevent FFT from being applied.

The first order probability density function of simulated process is given by (Yang 1973)

$$p(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\prod_{n=0}^{N-1} J_0(\theta\sqrt{2}A_n)e^{-i\theta x}d\theta \tag{2.24}$$

where J_0 is the zero-order Bessel function of the first kind. If the area under target spectrum is equally discretized such that

$$A_n^2 = 2S(\omega_n)\Delta\omega_n = \sigma^2/N \tag{2.25}$$

where σ is the standard deviation of target process, Eq. (24) will take the form

$$p(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\prod_{n=0}^{N-1} J_0^N(\theta\sqrt{2/N})e^{-i\theta x}d\theta \tag{2.26}$$

which is a function of N. It can be shown from Eq. (21) through Eq. (26) that the Gaussianness and ergodicity can be achieved when $N \to \infty$ (Shinozuka and Deodatis 1991). Some other essential properties of Eq (18) are listed as follows.

- The condition set in Eq (20) is necessary to guarantee that the temporal average and the temporal autocorrelation of any sample $x(t)$ are identical to those of target process.

- Under the condition of Eq (19), the sample process is periodic with period T_0, which is

$$T_0 = 2\pi / \Delta\omega \tag{2.27}$$

If Eq(22) is used instead, we have

$$T_0 = 4\pi / \Delta\omega \tag{2.28}$$

- According to the sampling theorem, Nyquist condition must be satisfied to prevent aliasing, namely

$$\Delta t \le 2\pi / 2\omega_u \tag{2.29}$$

- The value of sample process is bounded by

$$x(t) \le \sqrt{2}\sum_{n=0}^{N-1} A_n = \sqrt{2}\sum_{n=0}^{N-1} (2S(\omega_n)\Delta\omega)^{1/2} \tag{2.30}$$

It is readily to see that, the limit of spectral method is posed by Eq. (27) through Eq. (30). To relieve them, the period T_0 of generated sample must be long enough on the one hand to keep all the main statistical information, on the other hand, N must be big enough to allow a rational bound. Actually, even a relatively small N can offer a bound large enough in most of the cases.

The straightforward implementation of Eq. (18) is computationally costly. If the length of simulated process is M, it will involve the summation of $N \times M$ sinusoids. To improve the

efficiency, fast Fourier transform (FFT) can be used, which will reduce the operation to the order of $M\log_2 M$. The corresponding formulation of FFT is

$$x(p\Delta t) = \mathrm{Re}\left\{\sum_{n=0}^{M-1} B_n \exp\left[i\left(n\Delta\omega\right)\left(p\Delta t\right)\right]\right\}, \quad p = 0,1,\ldots,M-1 \tag{2.31}$$

where

$$B_n = \sqrt{2}A_n \exp(i\Phi_n), \quad n = 0,1,\ldots,M-1 \tag{2.32}$$

and

$$B_0 = B_n = 0, \quad N \le n \le M-1 \tag{2.33}$$

To apply FFT technique, the target spectrum must be discretized according to Eq. (19), and Nyquist condition must be satisfied. Since

$$M\Delta t = T_0 = 2\pi/\Delta\omega \tag{2.34}$$

Eq.(29) can be written as

$$M \ge 2N \tag{2.35}$$

This technique can also be extended to multidimensional cases (Shinoauka 1996). And other transforms like Hartley Transform may be used to increase the computational efficiency (Winterstein 1990).

2.6 Generation of non-Gaussian random process

Once a Gaussian stochastic series $x(t)$ is obtained, its non-normal counterpart g(t) can be generated by a non-linear transform

$$g(t) = Z[x(t)] \tag{2.36}$$

where Z is a non-linear function. Usually a zero memory non-linear function (ZMNL) is used for many of its favourable properties. If the cumulated density function (c.d.f) of g is G, Z can be simply chosen as

$$Z(x) = G^{-1}(F(x)) \tag{2.37}$$

where F is the c.d.f of normal distribution.

Given target spectrum of $g(t)$, if the spectrum of $x(t)$ is obtained, the spectral method in the previous section can be used again to generate $g(t)$. So it is in nature a kind of inverse problem. Assume $x(t)$ is a standard normal process without loss of generality, we can expand F in Hermite polynomials as

$$g(t) = \sum_{k=0}^{\infty} a_k H_k(t) \tag{2.38}$$

where

$$a_k = \frac{1}{k!}\int_{-\infty}^{\infty} g(t) H_k(t)\phi(t)dt \tag{2.39}$$

$H_k(t)$ is the k_{th} Hermite polynomial, and $\phi(t)$ is the p.d.f of standard normal distribution. So the target autocorrelation ρ of $g(t)$ can be written as (Madsen H. O. et. al. 1986)

$$\rho(\tau) = \sum_{k=1}^{\infty} b_k^2 [\rho_x(\tau)]^k \tag{2.40}$$

where ρ_x is the autocorrelation function of Gaussian input $x(t)$ and

$$b_k^2 = k! a_k^2 \tag{2.41}$$

Since $\rho(0) = \rho_x(0) = 1$, we have

$$\sum_{k=1}^{\infty} b_k^2 = 1 \tag{2.42}$$

Thus if $\rho(\tau)$ is given, $\rho_x(\tau)$ can be extracted from Eq. (40), and its corresponding spectrum $S_x(\omega)$ can be calculated.

However for specified target spectrum, not all solution of Eq. (40) is nonnegative definite unless the target spectrum and the distribution of $g(t)$ are compatible with each other. A necessary condition of compatibility for ZMNL by Whitt (1976) is

$$\rho(\tau) \geq \min E[g(t), g(t+\tau)] = \int_{-\infty}^{+\infty} y G^{-1}[1-G(y)]dG(y) \tag{2.43}$$

where ρ is the autocorrelation function of target process. For zero mean symmetric distribution this places no restriction on ρ, otherwise ρ is in general bigger than $-E[g^2]$. If the compatibility is not satisfied, S_x will be negative at some points, and Eq. (18) will fail. In this case a truncated form of S_x can be used as an substitute (Liu and Munson 1982), namely

$$S_x'(\omega) = \begin{cases} S_x(\omega), & S_x(\omega) \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{2.44}$$

Define the distortion of specified and the actual autocorrelation as

$$\varepsilon = \sum_{k=1}^{\infty} w(k)[\rho(k) - \rho_g(k)]^2 \tag{2.45}$$

where w_k is a even and nonnegative weighting function. If the corresponding autocorrelation ρ_x and ρ^g_x of normal field are close to each other, Eq. (45) can be expressed as

$$\varepsilon \approx \bar{\varepsilon} = \sum_{k=1}^{\infty} \overline{w}(k)[\rho_x(k) - \rho_x^g(k)]^2 \tag{2.46}$$

where

$$\begin{aligned} \overline{w}(k) &= w(k)[\frac{dT}{dy}(\rho_x(k))]^2 \\ T(y) &= \sum_{i=1}^{\infty} d_k^2 y^i \end{aligned} \tag{2.47}$$

According to Parseval's relation, if $\overline{w}(k)$ is constant we have

$$\bar{\varepsilon} = \frac{\overline{w}}{2\pi}\int_{-\pi}^{\pi}[S_x(\omega) - S_x^g(\omega)]^2\, d\omega \tag{2.48}$$

It is readily seen that S'_x defined in Eq. (44) will make $\bar{\varepsilon}$ minimum.

As in other non-linear problems, iteration method is feasible in the generation of non-Gaussian field. Typical of them is that by Yamazaki and Shinozuka (1988). However according to the author's experience, its convergence can not be ensured for want of solid mathematical support. Liu and Munson (1982) also suggested an algorithm to deal with the case where $\overline{w}(k)$ is not constant. Because the achievement may not deserve the extra efforts for large sample size Eq. (44) will be used in succeeding calculation. Smallwood (1997) gave a review of non-Gaussian field generation based on techniques other than spectral method.

In effect, if Z in Eq.(36) is ZMNL, the spectrum of $x(t)$ will not change significantly after transformation. In the sense of first order approximation, we might as well use the target spectrum itself to generate $x(t)$. The efficiency of this method can be shown in example 3.

2.7 Examples

2.7.1 Example 1

Five random number generators are tested in this example using the statistics in Section 3. Their parameters are listed in Table 2.1. Generator Rand1 was first proposed by Lewis, Goodman, and Miller in 1969. It has passed all the new-coined theoretical tests and gained a very good credit in application. After thorough study, Park and Miller (1988) recommended it as a good minimum standard against which the other generator can be compared. Rand2 is its shuffling version. Rand3 is a combined generator proposed by L'ecuyer (1988). It includes two LCGs and needs two sets of parameters. It is famous for the long period, which is $(m_1\text{-}1)(m_2\text{-}1)$. Rand4 is the shuffling from of Rand3. Rand5 is a shift register generator from IMSL (International Mathematics and Statistics Library). The results of test are listed in Table 2.2.

Obviously all the generators in Table 2.1 have passed the designed tests. However, the statistical properties are dependent on the seed. Furthermore, it should be noted that shuffling would not necessarily increase the randomness. It is successful for Lewis' scheme but it makes things a little worse for L'ecuyer's scheme. And the extent to which it will improve the randomness depends on the seed and the size and initial values of shuffling array. So it is highly recommended that we should always carry out necessary tests for the generator to be used, even if it belongs to a reliable commercial numerical library. The tests carried out here are fit for our purpose. Because Rand2 behaves a little better in the present tests, it will be used through the following research. The empirical distribution of random number from Rand2 is shown in Fig. 2.1, the coefficient of correlation is shown in Fig. 2.2. Here the u values are statistics from the generated samples. According to Eq. (7) and Eq. (10) our null hypothesis here is they all belong to the standard normal distribution. Given significant level $\alpha = 0.1$, we have fractile $F_1\text{-}\alpha_{/2} = 1.645$. The hypothesis is rejected when $|u| > 1.645$. Similarly in Chi-square test the fractile is $F_1\text{-}\alpha = 5128$. And the uniformity test will fail if $\chi^2 > 5128$.

Table 2.1 Parameters of random number generators

Name	a	q	r	M	Period	Seed	Shuffling
Rand1	16807	127773	2836	2147483647	2147483646	1	No
Rand2	16807	127773	2836	2147483647	2147483646	1	Yes
Rand3	40014	53668	12211	2147483563	2.3×10^{18}	1	No
	40692	52774	3791	2147483399		12345	
Rand4	40014	53668	12211	2147483563	2.3×10^{18}	1	Yes
	40692	52774	3791	2147483399		12345	
Rand5	-	-	-	2147483647	2147483646	1	No

Table 2.2 Statistical tests of random number generator

Name	$\bar{x}$	u_1	$\bar{x^2}$	u_2	s^2	u_3	χ^2	k_{50}
Rand1	0.4999	-0.04605	0.3332	-0.3101	0.8326	-1.062	4936	4
Rand2	0.4999	-0.05348	03332	-0.3177	0.8326	-1.064	4935	1
Rand3	0.4996	-1.492	0.3329	-1.491	0.8332	-0.1850	5025	2
Rand4	0.5000	0.02310	0.3333	-0.1284	0.8329	-0.6028	4971	3
Rand5	0.4997	-1.012	0.3330	-1.102	0.8329	-0.4893	5066	6

Note: $N = 1048576$, Significance level: $\alpha = 0.1$, $F_{1\text{-}\alpha/2} = 1.645$for u, $F_{1\text{-}\alpha} = 5128$ for χ^2 ($k = 5000$), k_{50} is the time of denial in the independence test for $\tau = 1(1)50$.

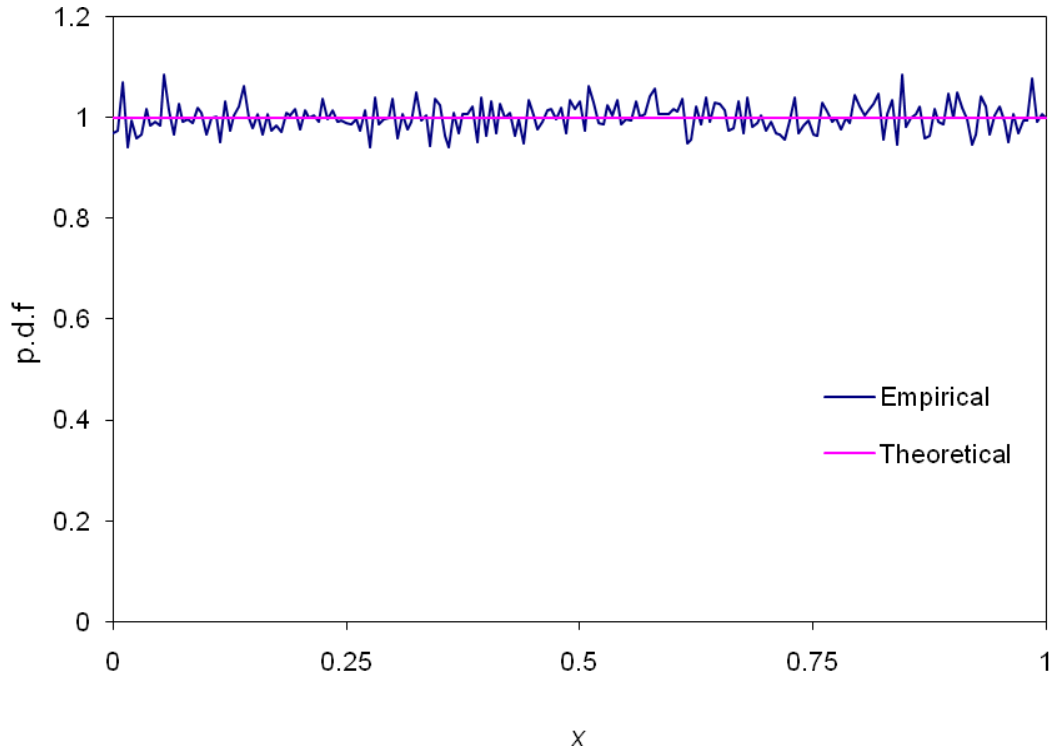


Figure 2.1 Empirical distribution of sample random series by Rand2

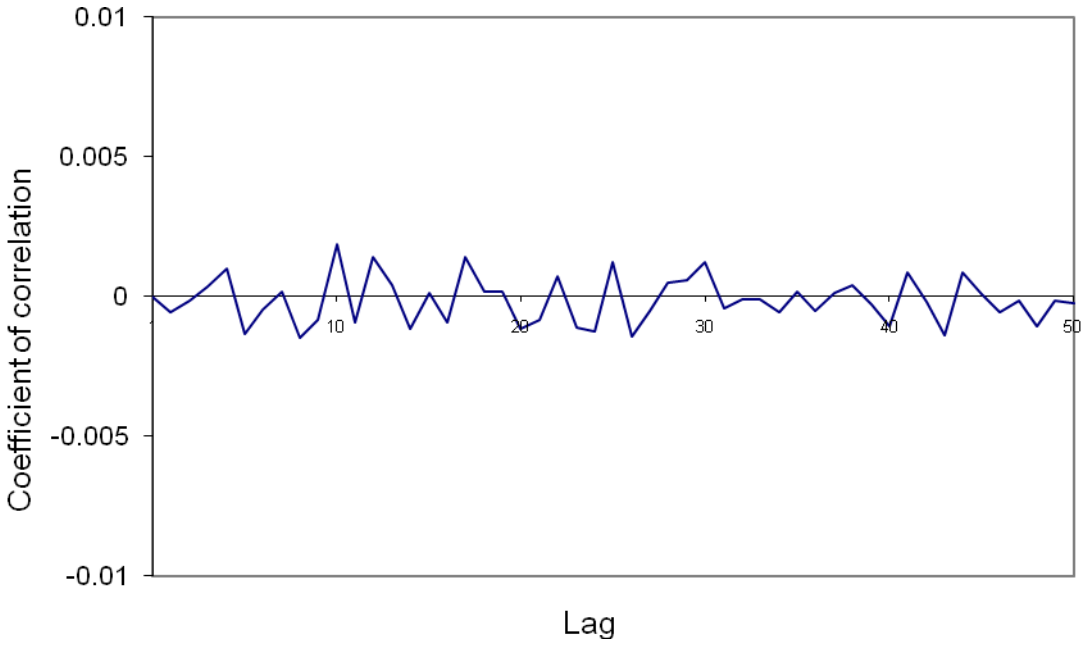


Figure 2.2 Coefficient of correlation of sample random series by Rand2

2.7.2 Example 2

A Gaussian random process with zero mean is generated by spectrum method in this example. The target spectrum density function is

$$S(\omega) = \frac{1}{4}\sigma^2 b^3 \omega^2 e^{-b|\omega|} \tag{2.49}$$

where σ^2 is the standard deviation of random field, b is correlation distance. According to Wiener-Khinchine relation, the autocorrelation function is

$$R(\tau) = \sigma^2 \frac{b^4(b^2 - 3\tau^2)}{(b^2 + \tau^2)^3} \tag{2.50}$$

This spectrum is also used by Shinozuka (1991). The parameters of target random field are listed in Table 2.3 for b = 1.

Table 2.3 Parameters of target field in Example 2.2

Parameter	Definition (Gaussian process)	Value
Deviation, σ	$(m_0)^{1/2}$	1.0
Zero crossing rate, ν_0	$(m_2 / m_0)^{1/2} / 2\pi$	0.5513
Rate of local maxima, n_0	$(m_4 / m_2)^{1/2} / 2\pi$	0.8717
Irregularity factor, α	ν_0 / n_0	0.6325
Spectral width, ε	$(1 - \alpha^2)^{1/2}$	0.7746

Note: m_k is the k_{th} moment of double-sided spectral density function $2\int_0^\infty \omega^k S(\omega) d\omega$

To generate the required field we chose

$$T_0 = 32768\,\text{sec}, \text{and } \omega_u = 4\pi \tag{2.51}$$

According to Eq.(19) and (27) the number of frequency step is

$$N = 65536 \tag{2.52}$$

This is more than enough to ensure Gaussianness and ergodicity (see Table 1 in [10]). According to Eq. (30), the upper bound of simulated process is

$$x(t) \le 0.2849 \times 10^3 \tag{2.53}$$

which can be regarded as a infinitive bound considering the unit deviation. To satisfy Nyquist condition set by Eq. (35), we chose

$$M = 2N = 131072 \tag{2.54}$$

with time step

$$\Delta t = 0.25\,\text{sec} \tag{2.55}$$

The simulated random process is shown in Fig. 2.3. Its spectrum density and autocorrelation function are shown in Fig. 2.4. Figure 2.5 gives the empirical distribution of sampling points. It is seen clearly that the precision of spectra method is very high. However the computation time is less than one second on a PC with only Pentium 133 processor.

As is shown in Fig. 2.3, the variation of simulated process with present time step is very severe, which may result from imperfection of perks and troughs. To study this imperfection a series of samples are generated with the same period but reduced time step. The estimates of zero-crossing rate and peak rate are calculated for each sample and listed in Table 2.4, along with temporal moments up to the fourth order.

Table 2.4 Statistics of simulated Gaussian field in Example 2

Time step (Sec)	Mean	Mean square	Skewness	Kurtosis	Peak rate	Zero-crossing rate
¼	0.4244×10^{-18}	0.9997	$-.7488\times10^{-2}$	3.0272	0.7733	0.5271
1/8	$-.5485\times10^{-19}$	0.9997	$-.7602\times10^{-2}$	3.0276	0.8394	0.5446
1/16	0.2433×10^{-18}	0.9997	$-.7602\times10^{-2}$	3.0276	0.8567	0.5497
1/32	$-.6072\times10^{-19}$	0.9997	$-.7602\times10^{-2}$	3.0276	0.8610	0.5506
1/64	$-.2429\times10^{-17}$	0.9997	$-.7602\times10^{-2}$	3.0276	0.8620	0.5510
1/128	0.2939×10^{-18}	0.9997	$-.7602\times10^{-2}$	3.0276	0.8624	0.5510
1/256	$-.5895\times10^{-17}$	0.9997	$-.7602\times10^{-2}$	3.0276	0.8624	0.5511

It can be noticed that all the statistics in Table 2.4 tend to converge when time step decreases. The convergent points are very close to theoretical results. In addition it is found that the convergence speed of peak rate and zero-crossing rate is lower than that of moment estimates. Usually, the effect of peak imperfection can be ignored and a relatively bigger time step can be used. However, in fatigue analysis, it may cause significant error in the counting of rain flow ranges. So special attention has to be paid in this case. The simulated time series with step Δt = 1/32 is shown in Fig. 2.6 with good continuity.

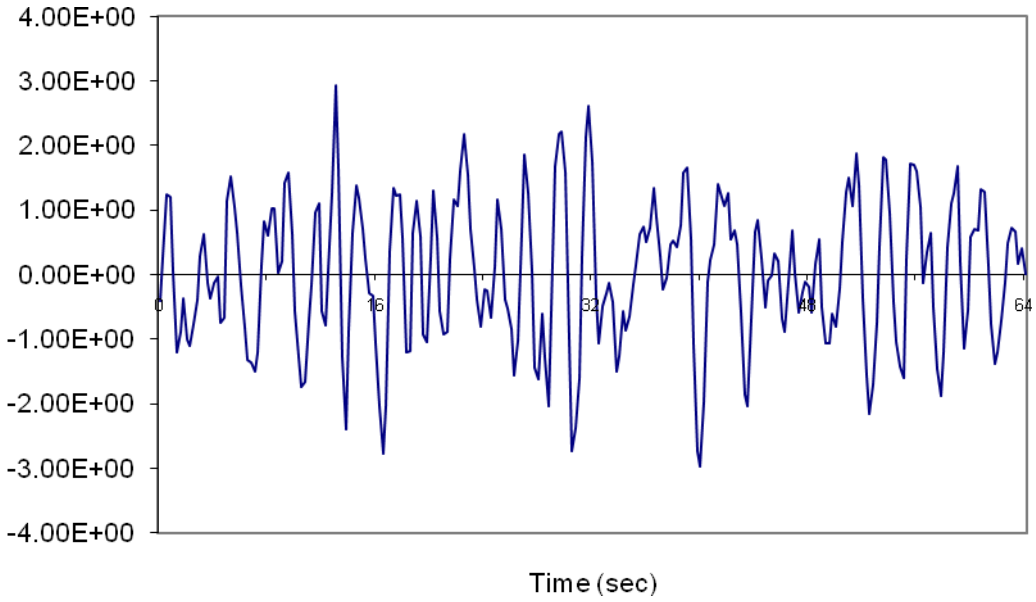


Figure 2.3 Time series of simulated Gaussian process (Δt =1/4sec)

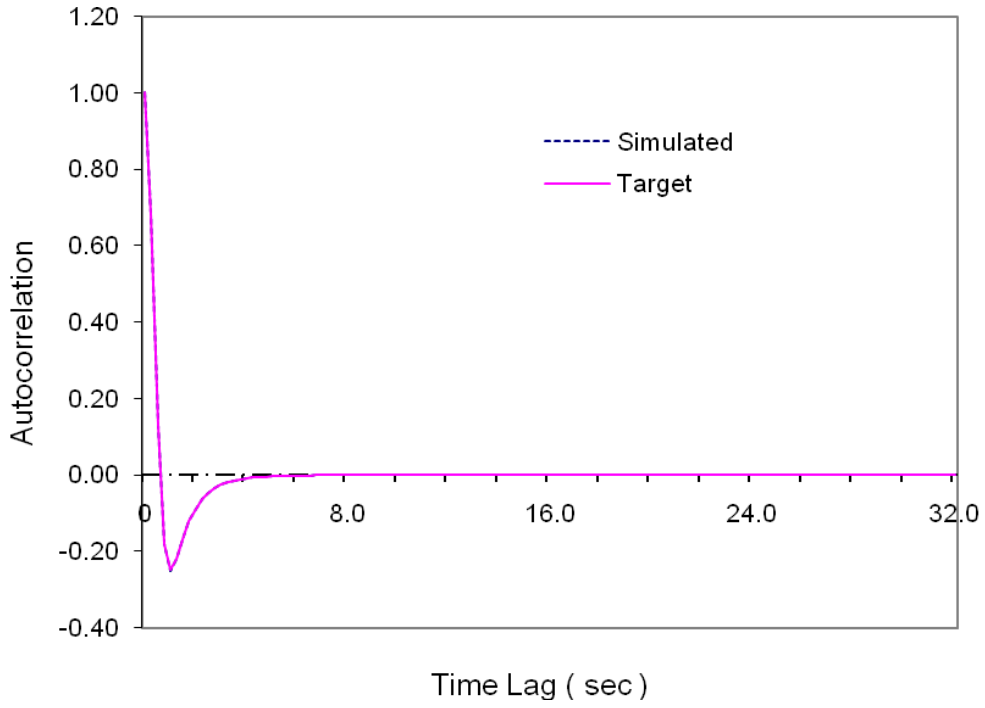


Figure 2.4(a) Temporal autocorrelation of simulated Gaussian process (Δt=1/4sec)

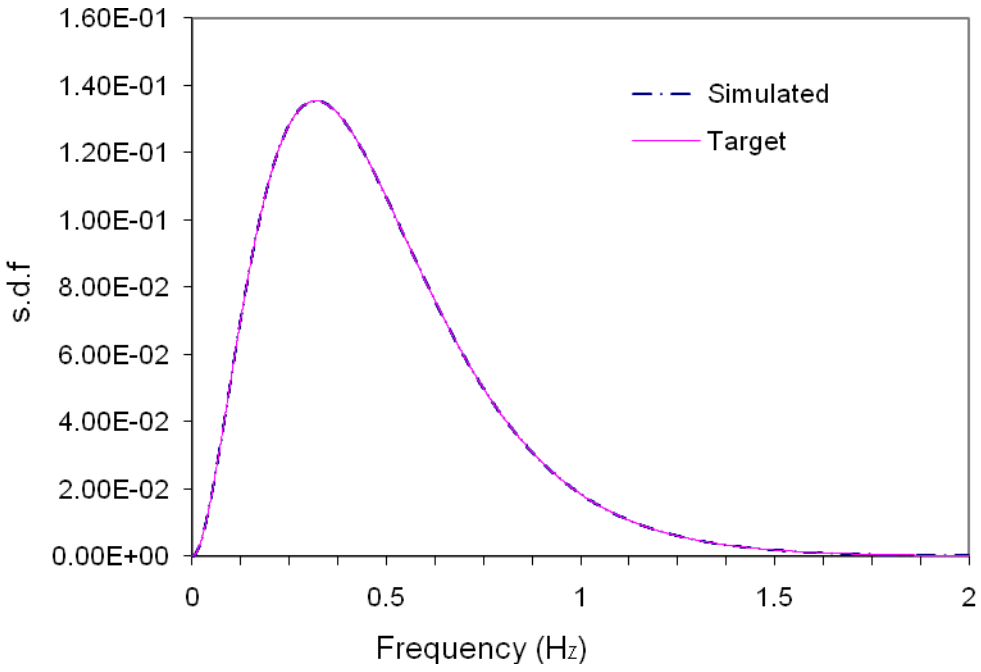


Figure 2.4(b) Spectral density of simulated Gaussian process (Δt=1/4sec)

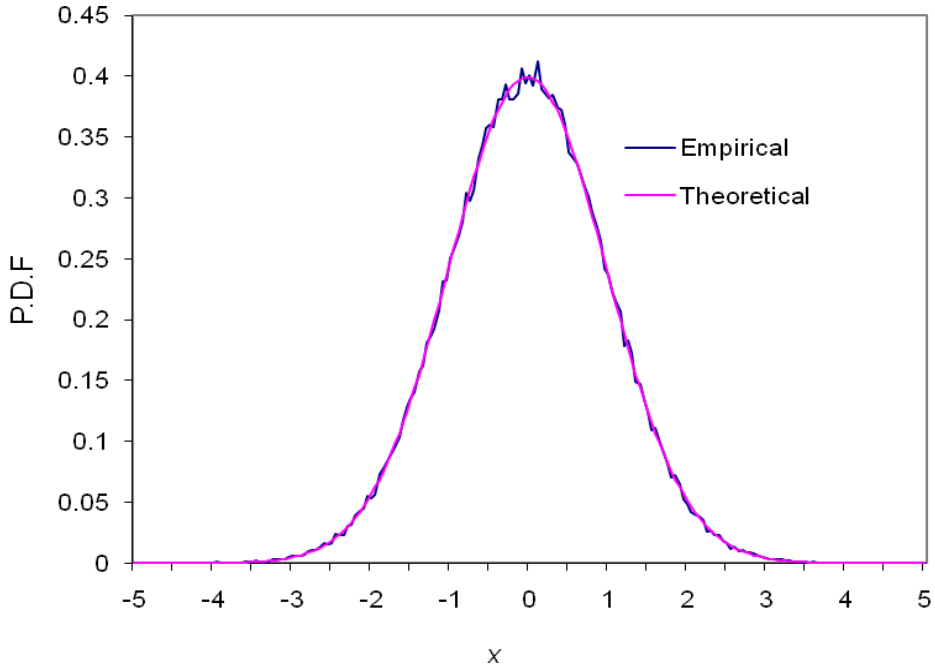


Figure 2.5 Empirical distribution of simulated Gaussian process

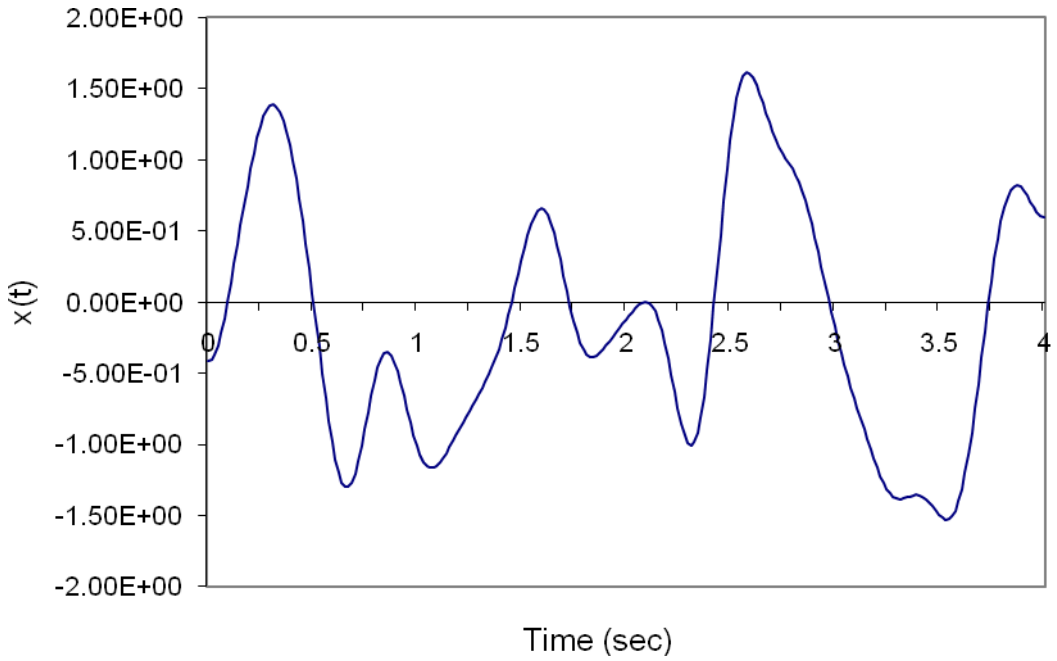


Figure 2.6 Time series of simulated Gaussian field (Δt =1/64sec)

2.7.3 Example 3

In this example a non-Gaussian process will be generated based on the Gaussian process generator in Example 2. The ZMNL transformation is chosen as

$$g = F(x) = \frac{\sinh(\frac{x-\gamma}{\delta}) - \mu_1'}{\sqrt{\mu_2}} \tag{2.56}$$

where γ and δ are shape parameters, and μ'_1 and μ_2 can be calculated by

$$\begin{aligned} \mu_1' &= -\sqrt{\omega}\sinh(\Omega) \\ \mu_2 &= \frac{1}{2}(\omega - 1)(\omega\cosh(2\Omega) + 1) \end{aligned} \tag{2.57}$$

where

$$\begin{aligned} \omega &= \exp(\delta^{-2}) \\ \Omega &= \frac{\gamma}{\delta} \end{aligned} \tag{2.58}$$

The random variable g defined in Eq. (52) is zero mean and univariate. It is a standardised version of S_U family (Johnson 1972). Let $\gamma = 5.0$ and $\delta = 2.0$, the skewness and kurtosis of g are 1.731 and 8.801 respectively, which shows strong non-linearity.

The random process of $g(t)$ is generated by Eq. (56) with the same target spectral density as in Example 2. In generator RFSU1, the spectral density function used to generate Gaussian input is calculated by Eq. (40) and truncated by Eq. (44). While in generator RFSU2 the target spectral density is used directly to generate Gaussian input. They are performed by choosing T_0 = 32768sec, Δt = 1/64sec, and N = 65536. The autocorrelation and spectral density of their outputs are shown in Fig. 2.7 and Fig. 2.8. The time history of both generators is given by Fig. 2.9 and the first four sample moments are listed in Table 2.5.

As can be seen, the time series by generator with optimised distortion preserves the statistical characteristics of target process quite well. While the distortion of generator

without optimisation is also pretty tolerable. It is reasonable to believe that the distortion in both generators will decrease when the non-linearity of Eq. (36) reduces.

Table 2.5 Statistics of simulated S_U field in Example 2.3

Moments of sample	RFSU1	RFSU2
Mean	$-.1422\times10^{-2}$	$-.358410^{-3}$
Deviation	1.015	1.009
Skewness	-1.832	-1.800
Kurtosis	9.469	9.198

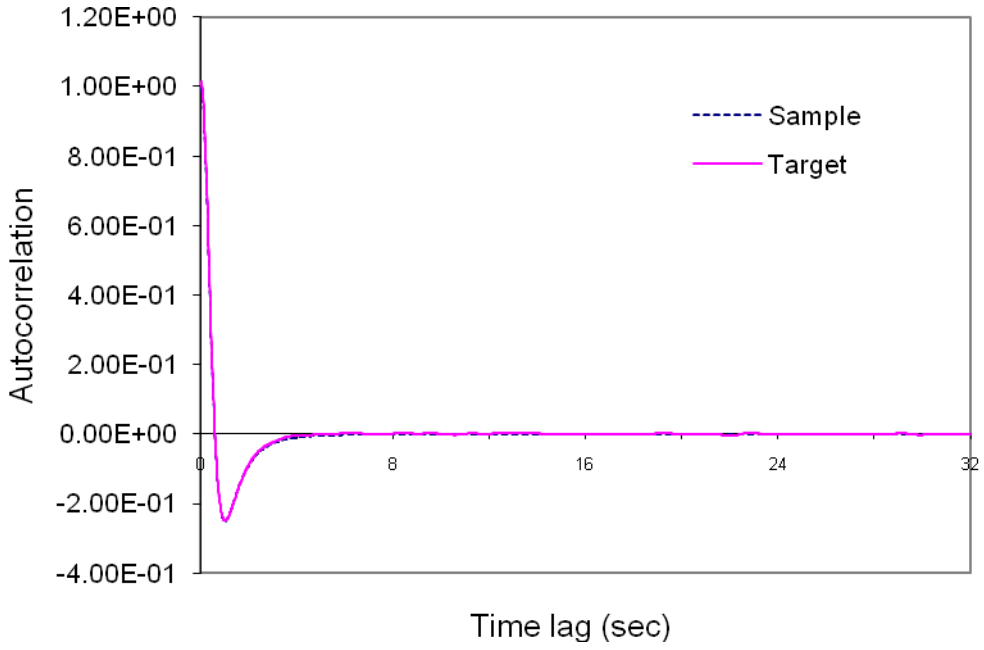


Figure 2.7(a) Autocorrelation function of simulated S_U process by RFSU1

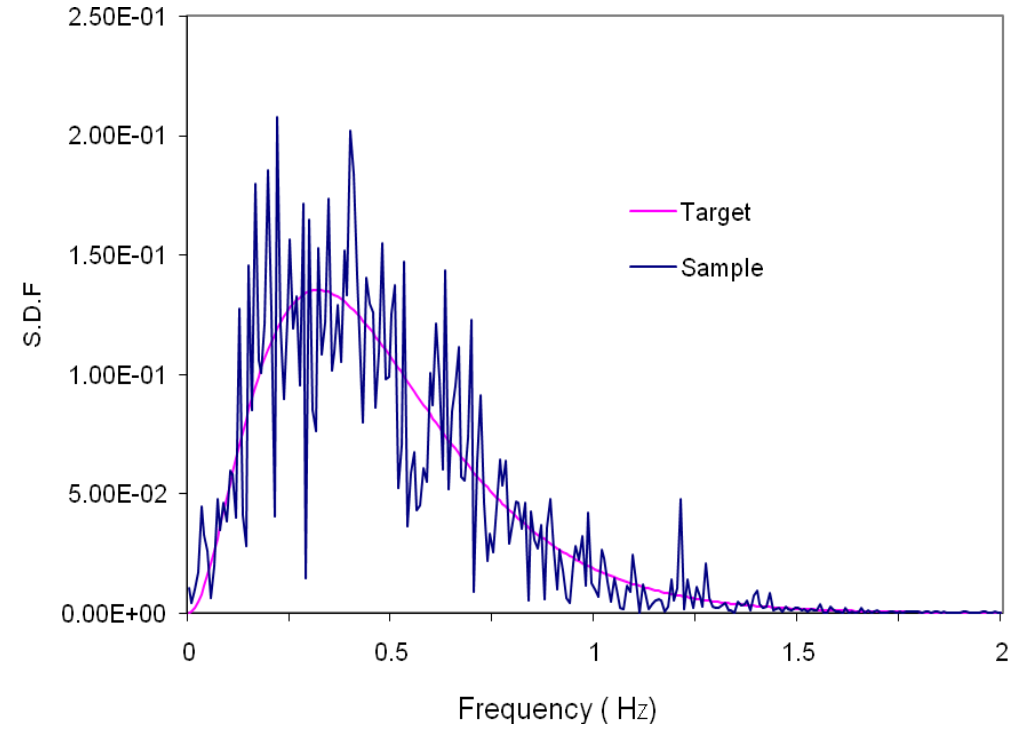


Figure 2.7(b) Spectral density function of simulated S_U process by RFSU1

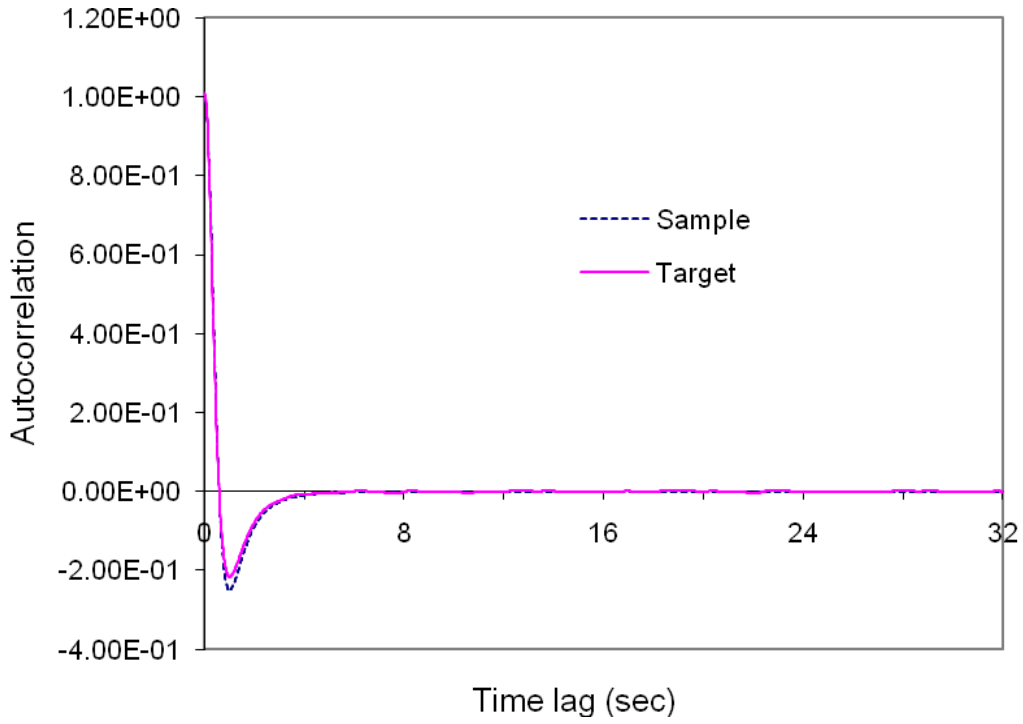


Figure 2.8(a) Autocorrelation function of simulated S_U field by RFSU2

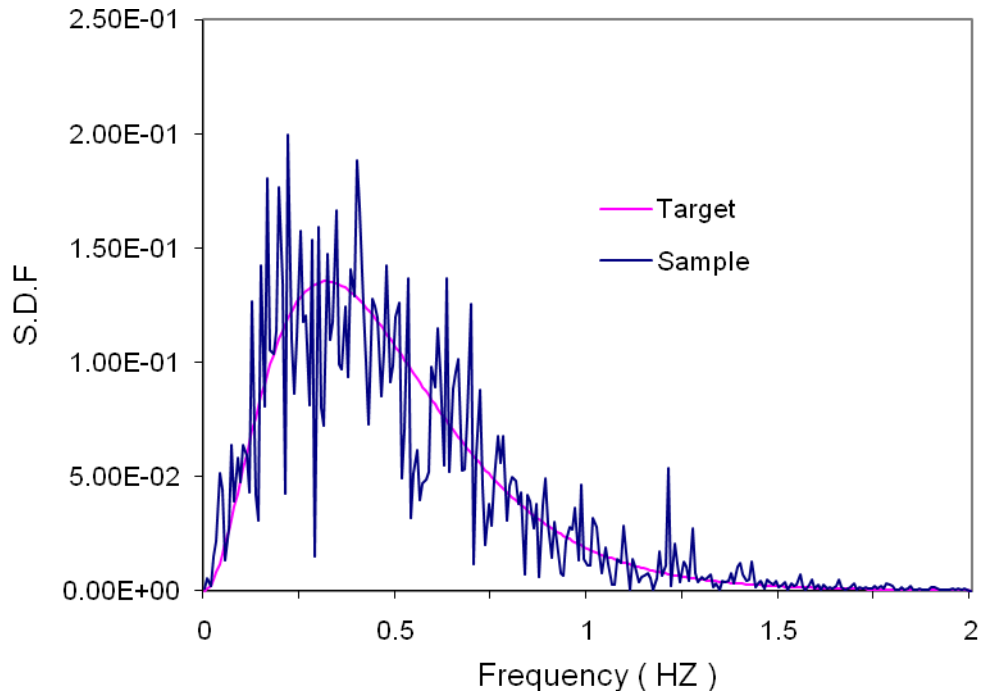


Figure 2.8(b) Spectral density function of simulated S_U process by RFSU2

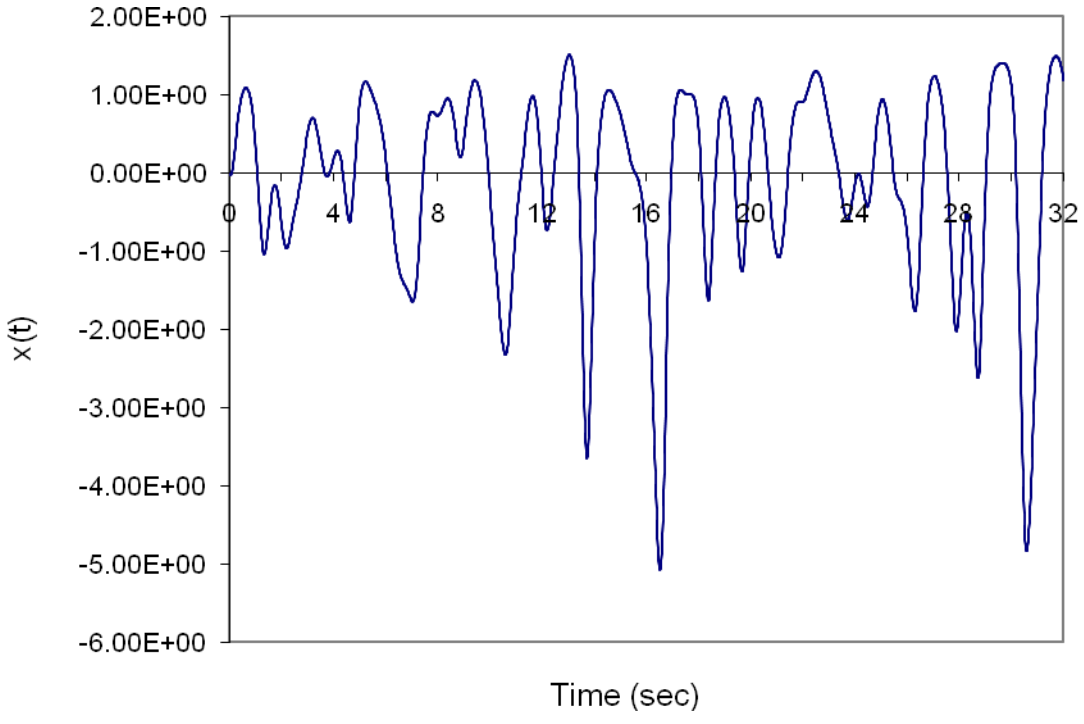


Figure 2.9(a) Time series of S_U field by RFSU1

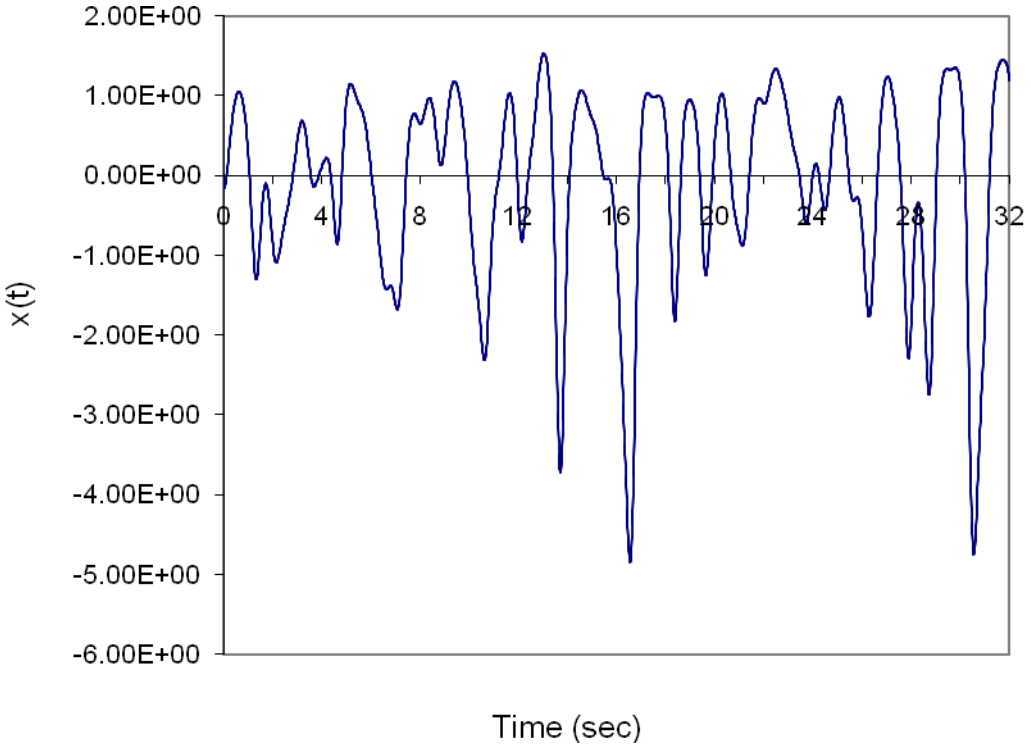


Figure 2.9(b) Time series of S_U field by RFSU2

2.8 Conclusions

- Parameters play an important part in random number generation. For more reliable results in specific area, necessary statistical tests have to be carried out even for those renowned generators.

- The spectral method aided by FFT shows both high precision and high efficiency in random process generation. It is ideal for large-scale simulation.

- Nyquist condition can not ensure the peak and trough perfection in simulated field. For fatigue analysis using rain flow counting, a reasonably small time step should be chosen.

- The compatibility of given spectrum and distribution is usually not satisfied. So the generated non-Gaussian field is more or less distorted. Though optimum method can be used to reduce the distortion to the minimum under certain criteria, it involves extra computational cost. So we might as well do without it especially in mild nonlinearity problem, as is often the case.

References

1. Liu, B. & Munson, D. C. (1982) Generation of a random sequence having a jointly specified marginal distribution and autocovariance. *IEEE Transaction on Acoustics, Speech, and Signal Processing*, 30(6), 973-983.

2. Johnson, N. L. & Kotz, S. (1970). *Distributions in statistics: continuous univariate distribution*. New York: Houghton Mifflin.

3. Knuth, D. E. (1981). *The art of computer programming* (Vol. 2). London: Addison-Wesley.

4. L'ecuyer, P. (1988). Efficient and portable combined random number generators. *Communications of ACM*, 31(6), 742-749.

5. Madsen, H. O., Krenk, S. and Lind N. C. (1986). *Methods of structural safety*. New Jersey: Prientice-Hall.

6. Newman, M. E. J. & Barkema G. T. (1999) *Monte Carlo methods in statistical Physics*, Oxford: Clarendon Press.

7. Papoulis, A. (1965). *Probability, random variables, and stochastic process*. New York: McGraw-Hill.

8. Park, S. T. & Miller, K. W. (1988). Random number generations: good ones are hard to find. *Communications of the ACM* , 31(10), 1192-1210.

9. Rubinstein, R. Y. (1981). *Simulation and the Monte Carlo methods*. New York: John Wiley & Sons.

10. Shinozuka, M., & Deodatis, G. (1991). Simulation of stochastic process by spectral representation. *Applied Mechanics Reviews*, 44(4), 191-204.

11. Shinozuka, M. & Deodatis, G. (1996) Simulation of multi-dimensional Gaussian fields by spectral representation. *Applied Mechanics Reviews*, 49(1), 29-53.

12. Samllwood, D. O. (1997). Generation of stationary non-Gaussian time histories with a specified cross-spectral density. *Shock and Vibration*, 4(5-6), 361-377.

13. Spanos, P., & Zeldin, B. (1998). Monte Carlo treatment of random fields: a broad perspective. *Applied Mechanics Reviews*, 51(3), 219-237.

14. Vanmarcke, E. (1983). *Random fields: analysis and synthesis*. Massachusetts: The MIT Press.

15. Yamazaki, F., & Shinozuka, M. (1988). Digital generation of non-Gaussian stochastic fields. *Journal of Engineering Mechanics*, 114(7), 1183-1197.

Chapter 3 Second-Moment Theory of Reliability Analysis

3.1 Introduction

Structural reliability analysis requires more information than in deterministic analysis. Theoretically, if the p.d.f. or the joint p.d.f. of state variables is available, the corresponding failure probability can be evaluated directly through a multidimensional integration. However accurate and efficient numerical integration is often an insurmountable task in practice for given distributions, to say nothing that there are many cases where the complete probabilistic information is unknown for want of sufficient data or knowledge. As a result, normal distribution (the bell curve) is often resorted to as a practical alternative. Since the first two moments, namely, the mean value, the variance and the covariance, are sufficient to describe a normal random vector, and can be evaluated easily from the available information, it is often preferable to perform reliability analysis in the normal space. Based on this very concept, the second-moment reliability theory is established.

A variety of second-moment based researches were carried out before the sixties. The cornerstone was laid by Freudenthal (1956) who used complete probability models. However, it is the work of Cornell (1969) that heralded popular acceptance of second moment concept. Later, among many other researchers, Shinozuka (1983) presented a brand new interpretation to the theory. To date, second-moment approaches have become so popular that it always takes an important place in the text books concerning structural safety. Typical of them are those by Ang and Tang (1984), Madsen, Krenk and Lind (1986), Ditlevsen and Madsen (1996), Zhao(1996), and Melcher (1999). With reference to the works mentioned above, the chapter is presented to discuss in detail the second-moment theory in a light of geometry and optimization. A new set of formulations are developed in this chapter for general reliability cases, viz. when the basic variables are

non-normal and correlated. Both first order reliability method (FORM) and second order reliability method (SORM) are given due consideration. As shown in the chapters that follow, they serve as a primary technique in RSM. All the algorithms developed are tested on delicately designed examples. It is shown that both high efficiency and precision can be achieved.

3.2 Geometric measure of reliability

It is now a common sense that geometry and algebra are closely related to each other. In fact, this kinship is the foundation on which the building of modern applied mathematics is constructed. In some sense, almost all the problems in practice may finally be converted equivalently to a problem of minimum distance in certain space, for example, the well known light traveling problem in Euclidian space, the least square problem in Hilbert space, and the variational problem in Soblevian space. Little wonder this also holds true of reliability analysis.

To illustrate this point, let us consider a system with limit state function

$$G(\boldsymbol{X}) = G(X_1, X_2, \ldots, X_n) \tag{3.1}$$

where $\boldsymbol{X}$ is the vector of state variables, which are assumed to be uncorrelated normal ones without loss of generality (otherwise some transformations can be used to meet the assumption). Geometrically, the limit state equation $G(\boldsymbol{X}) = 0$ is an n-dimensional surface in random space. It is termed failure surface, with one side $G(\boldsymbol{X}) > 0$ defining the safe state and the other side $G(\boldsymbol{X}) < 0$ the failure state. It follows that the failure probability is

$$P_f = P\{G(\mathbf{X}) \le 0\} \tag{3.2}$$

if $M = G(\boldsymbol{X})$, the safety margin, happens to be a normal variable then

$$P_f = P\{M \le 0\} = F_M(0) = \Phi(-\mu_M / \sigma_M) \tag{3.3}$$

where F_M is the c.d.f. of M, Φ is the c.d.f. of standard normal variable, and μ_M and σ_M are the mean value and the standard deviation of M. Defining

$$\beta = \mu_z / \sigma_z = \frac{E[G(\mathbf{X})]}{D[G(\mathbf{X})]} \tag{3.4}$$

we obtain

$$P_f = \Phi(-\beta) \tag{3.5}$$

where β is termed reliability index or safety index. The definition in Eq. (4) is first given by Cornell (1969). It is the distance of location $E[G(\boldsymbol{X})]$ to the failure surface $M = 0$ in the unit of $D[G(\boldsymbol{X})]$.

As is always the case, $G(\boldsymbol{X})$ may not be a linear function of $\boldsymbol{X}$; that is, M is not necessarily normal. For convenience, a hyper-plane $R(\boldsymbol{X}) = 0$ tangent to the failure surface may be adopted as an approximate limit state by expanding $G(\boldsymbol{X})$ in Taylor series at a point $\boldsymbol{X}_0$ on the actual failure surface and retaining the first order items only. Considering that $G(\boldsymbol{X}_0) = 0$, this yields

$$R(\mathbf{X}) = \mathbf{G}_0^T \left(\mathbf{X} - \mathbf{X}_0 \right) \tag{3.6}$$

where

$$\mathbf{G} = \nabla G = \left(\frac{\partial G}{\partial X_1}, \frac{\partial G}{\partial X_2}, \ldots, \frac{\partial G}{\partial X_n} \right)^T \tag{3.7}$$

Introducing the standardized vector

$$\overline{\mathbf{X}} = \mathbf{s}^{-1} \left(\mathbf{X} - \mathbf{m} \right) \tag{3.8}$$

with

$$\mathbf{m} = E[\mathbf{X}] \tag{3.9}$$

and

$$\mathbf{s} = diag\left(\sigma_{X_1}, \sigma_{X_2}, \cdots, \sigma_{X_n}\right) \tag{3.10}$$

We have

$$X = s\overline{X} + m \tag{3.11}$$

Substituting Eq. (11) into (6) gives a linear failure surface in standardized form

$$\overline{R}(\overline{\mathbf{X}}) = \mathbf{G}_0^T \mathbf{s}\overline{\mathbf{X}} + \mathbf{G}_0^T(\mathbf{m} - \mathbf{X}_0) = 0 \tag{3.12}$$

and the corresponding Cornell safety index is, according to (3.4),

$$\beta = \frac{E[\overline{R}(\overline{\mathbf{X}})]}{D[\overline{R}(\overline{\mathbf{X}})]} = \frac{\mathbf{G}_0^T(\mathbf{m} - \mathbf{X}_0)}{\left(\mathbf{G}_0^T \mathbf{s}^2 \mathbf{G}_0\right)^{1/2}} \tag{3.13}$$

Let us transform (3.12) into

$$-\frac{sign\left[\mathbf{G}_0^T(\mathbf{m} - \mathbf{X}_0)\right]}{\left(\mathbf{G}_0^T \mathbf{s}^2 \mathbf{G}_0\right)^{1/2}} \mathbf{G}_0^T \mathbf{s}\overline{\mathbf{X}} - \frac{\left|\mathbf{G}_0^T(\mathbf{m} - \mathbf{X}_0)\right|}{\left(\mathbf{G}_0^T \mathbf{s}^2 \mathbf{G}_0\right)^{1/2}} = 0 \tag{3.14}$$

Comparing with the standard form of a linear surface

$$N_0^T \overline{X} - d = 0 \tag{3.15}$$

where $\boldsymbol{N}_0$ is the unit vector normal to the surface, and d (≥ 0) is the distance from the origin to the surface, we obtain

$$d = \frac{\left|\mathbf{G}_0^T(\mathbf{m} - \mathbf{X}_0)\right|}{\left(\mathbf{G}_0^T \mathbf{s}^2 \mathbf{G}_0\right)^{1/2}} = |\beta| \tag{3.16}$$

and

$$\mathbf{N}_0^T = -\frac{sign\left[\mathbf{G}_0^T\left(\mathbf{m}-\mathbf{X}_0\right)\right]}{\left(\mathbf{G}_0^T\mathbf{s}^2\mathbf{G}_0\right)^{1/2}}\mathbf{G}_0^T\mathbf{s} \tag{3.17}$$

This proves that if $\boldsymbol{X}$ is an uncorrelated normal vector, the absolute value of safety index in the first order approximation is exactly the distance from the origin to one of the tangent planes of the actual failure surface in the standardized random space. Figure 3.1 gives a graphical interpretation of β for two-dimensional problem.

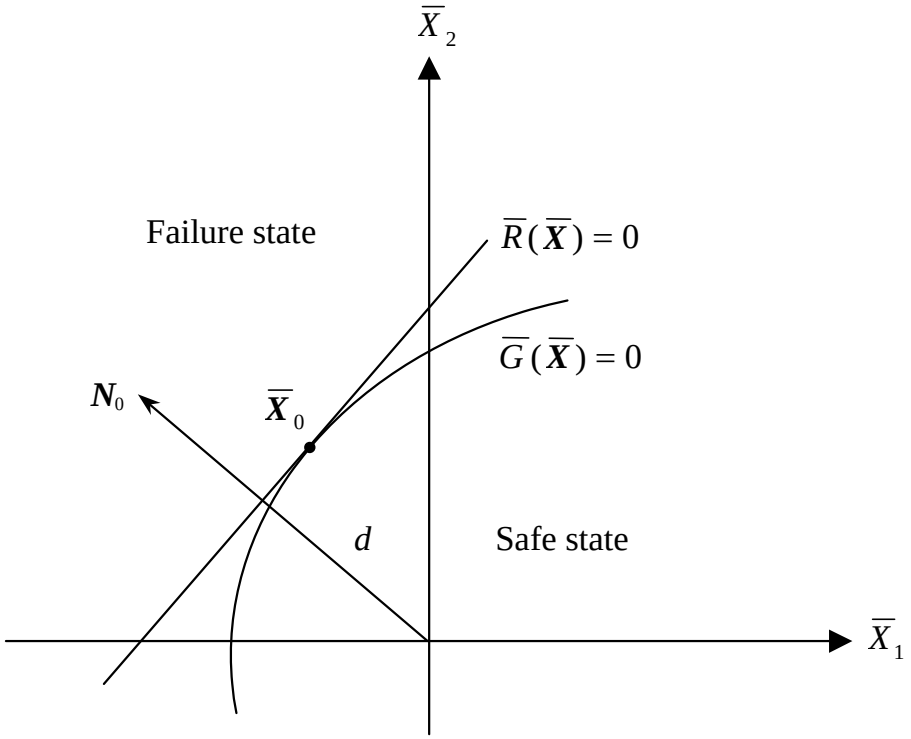


Figure 3.1 Linearization of limit state surface at arbitrary point

Certainly, the linearization of $G(\boldsymbol{X})$ will induce in P_f a truncation error. If $\overline{G}(\overline{\boldsymbol{X}})$ is concave to the origin and $\beta > 0$, the error can be expressed by

$$\begin{aligned}\Delta P_f &= \int_{\overline{G}(\overline{\mathbf{x}})\le 0} \bar{f}_{\overline{\mathbf{X}}}(\overline{\mathbf{x}})d\overline{\mathbf{x}} - \int_{\overline{R}(\overline{\mathbf{x}})\le 0} \bar{f}_{\overline{\mathbf{X}}}(\overline{\mathbf{x}})d\overline{\mathbf{x}} \\ &= P_f - \Phi(-d)\end{aligned} \tag{3.18}$$

where $\bar{f}_{\overline{\boldsymbol{X}}}(\overline{\boldsymbol{X}})$ is the joint c.d.f. of the standard normal variables. In order to reduce $|\Delta P_f|$, d should be decreased, hence, the least absolute error $|\Delta P_f|_{\min}$ will be achieved when d is the minimum distance from the origin to the failure surface $\overline{G}(\overline{\boldsymbol{X}}) = 0$. The

corresponding safety index β thus obtained is due to Hasofer and Lind (1974). It has been widely adopted as a practical measure of structural safety, and the relevant point $\overline{X}^*$ in the standardized random space and its counterpart $\boldsymbol{X}^*$ in the original space at which $G(\boldsymbol{X})$ is expanded are termed design point. According to Freudenthal (1956) an effective checking point in design should be the most probable failure point or a point of the maximum likelihood. Design point is such a point in standard space, about which $\bar{f}_{\overline{\boldsymbol{X}}}(\overline{\boldsymbol{X}})$ is symmetric. Likewise, the same conclusion can be drawn, when $\beta < 0$ or $\overline{G}(\overline{\boldsymbol{X}})$ is convex. For the present case, an interval estimation of P_f is given by Hasofer (1974), as follows

$$\Phi(-\beta) < P_f < 1-\chi_n^2(\beta^2) \tag{3.19}$$

where $\chi_n^2(\cdot)$ is the c.d.f. of the chi-square distribution with n degrees of freedom. This is a natural conclusion if we take the safe state as bounded between the space with hyper-plane $\overline{R}(\overline{\boldsymbol{X}}^*)=0$ and the hyper-sphere of radius β, while noticing the fact that the sum of squares of a set of uncorrelated standard normal variables has a chi-square distribution with n degrees of freedom. Here we assume there is only one design point and that the failure surface is concave to the origin. Figure 3.2 illustrates the implication of design point $\overline{X}^*$, the Hasofer and Lind safety index β, and the error ΔP_f for two-dimensional problem.

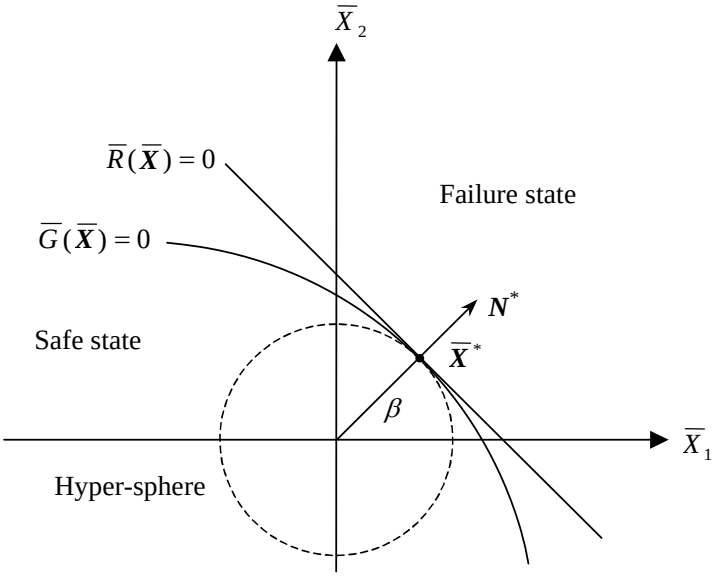


Figure 3.2 Hasofer and Lind reliability index

Unfortunately, Eq. (19) can hardly be used in practice, partly because it is too coarse to give an accurate estimate, partly because the curvature of $\overline{G}(\overline{X})$ is complicated in general. By "general" here we mean the failure surface may be convex and concave alternately around design point. Given a small error ε, there must exist an $\overline{X}^*$-centered hyper-sphere of radius r, which is divided by $\overline{G}(\overline{X}) = 0$ into failure domain $\overline{D}_f$ and safe domain $\overline{D}_s$(as shown in Fig. 3.3 for two-dimensional case), that is

$$\left| P_f - \int_{\overline{D}_f} \bar{f}_{\overline{\mathbf{X}}}(\overline{\mathbf{x}}) d\overline{\mathbf{x}} \right| \le \varepsilon \tag{3.20}$$

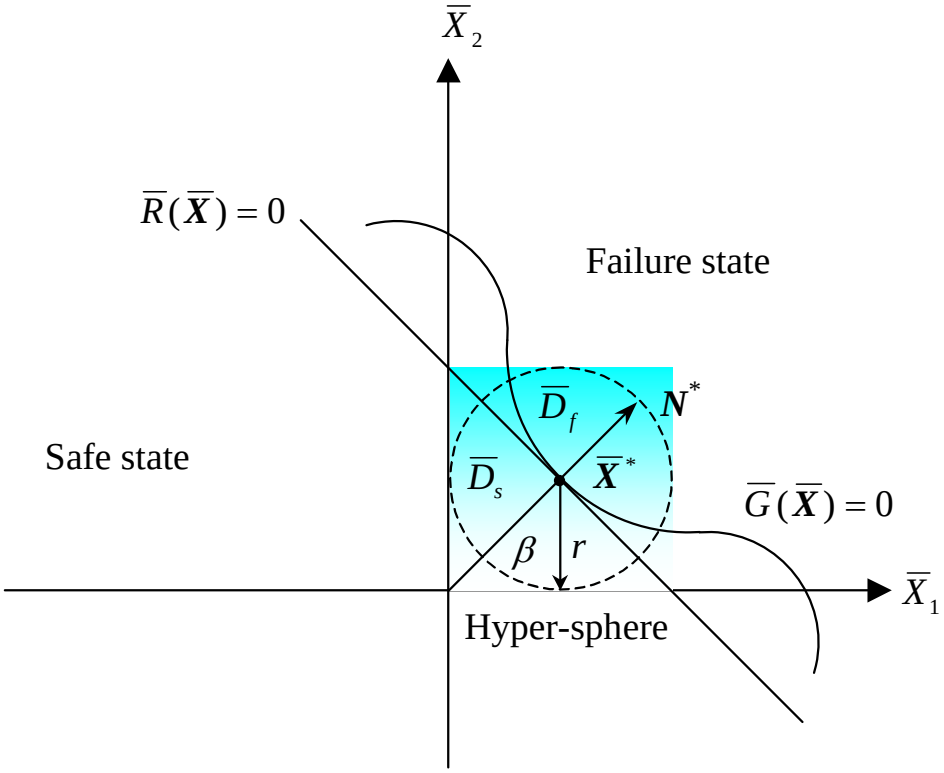


Figure 3.3 Illustration of general failure surface

The joint p.d.f. in standard normal space is

$$\begin{aligned} \bar{f}_{\overline{\mathbf{X}}}(\overline{\mathbf{x}}) &= \prod_{i=1}^{n} \phi(\overline{x}_i) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} \overline{x}_i^2 \right) \\ &= \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{d^2}{2} \right) \end{aligned} \tag{3.21}$$

Evidently, $\bar{f}_{\bar{X}}(\bar{x})$ decrease exponentially with *d*. Thus, a small *r* will satisfy (3.20). If $\overline{G}(\overline{X}) = 0$ is smooth enough, it will be either concave or convex approximately in such a small region. So our comments on $\overline{X}^*$ and β still hold true. This is very important to simplify the problem as in SORM or in Monte Carlo integration, which are usually performed in the vicinity of $\overline{X}^*$.

It should be noted that although Hasofer and Lind reliability index alleviates the ambiguity in linearization, it has no one-to-one relation with failure probability unless the limit state function in standard normal space happens to be a super-plane. For highly nonlinear problems, it is not a practical measure of reliability. This will be discussed in the final part of the chapter.

3.3 First-order formulation by Lagrangian method

As shown in the last section, the determination of safety index is equivalent to a problem of minimum distance in standard normal space, or rather, an optimization problem. This lends us a powerful means to make reliability analysis. In this section a simple first-order formulation by Lagrangian method is given. The more general formulations based on optimization theory will be developed in section 3.5.

Now let us consider the limit state function $G(\boldsymbol{X})$ in the standardized space of $\overline{\boldsymbol{X}}$ with covariance matrix $\overline{\boldsymbol{C}}$. Hereafter we name such a space correlated or dependent normal space. This slacks a little the requirements in section 3.2. An independent normal vector $\boldsymbol{U}$ can be constructed by appropriate linear transformation

$$\boldsymbol{U} = \boldsymbol{A}\overline{\boldsymbol{X}} \tag{3.22}$$

Accordingly, the limit state function will be transformed into $G_u(\boldsymbol{U})$. The transformation matrix $\boldsymbol{A}$ in Eq. (22) can be determined either by spectral decomposition (Vanmarcke 1983) or more simply by Cholesky decomposition (Ditlevsen 1981) since $\overline{\boldsymbol{C}}$ is usually positive definite. By spectral decomposition we have

$$\boldsymbol{A} = \boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{P} \tag{3.23}$$

where $\boldsymbol{P}$ is the orthogonal eigen-vectors of $\overline{\boldsymbol{C}}$, and $\boldsymbol{\Lambda}$ is a diagonal matrix of the corresponding eigen-values:

$$\boldsymbol{P}\overline{\boldsymbol{C}}\boldsymbol{P}^T = \boldsymbol{\Lambda} \tag{3.24}$$

It follows that

$$A\overline{C}A^T = I \tag{3.25}$$

and

$$\overline{C} = A^{-1}(A^{-1})^T \tag{3.26}$$

By Cholesky decomposition the correlation matrix can be written as

$$\overline{\boldsymbol{C}} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^T \tag{3.27}$$

where $\boldsymbol{L}$ is a lower triangular matrix with unit diagonal elements, and $\boldsymbol{D}$ is a diagonal matrix. Considering that

$$Cov(\boldsymbol{U}) = \boldsymbol{A}\overline{\boldsymbol{C}}\boldsymbol{A}^T = \boldsymbol{I} \tag{3.28}$$

we obtain

$$\boldsymbol{A} = \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{L}^{-1} \tag{3.29}$$

Again Eq. (26) holds. Now, the evaluation of safety index becomes such an optimization problem as minimizing distance

$$D = (\boldsymbol{U}^T\boldsymbol{U})^{1/2} = (\overline{\boldsymbol{X}}^T\boldsymbol{A}^T\boldsymbol{A}\overline{\boldsymbol{X}})^{1/2} = (\overline{\boldsymbol{X}}^T\overline{\boldsymbol{C}}^{-1}\overline{\boldsymbol{X}})^{1/2} \tag{3.30}$$

subject to the constraint

$$G_u(\boldsymbol{U}) = \overline{G}(\overline{\boldsymbol{X}}) = 0 \tag{3.31}$$

So we can construct a function L with multiplier λ in terms of $\overline{X}$

$$L = D + \lambda G_u(\boldsymbol{U}) = (\overline{\boldsymbol{X}}^T \overline{\boldsymbol{C}}^{-1} \overline{\boldsymbol{X}})^{1/2} + \lambda \overline{G}(\overline{\boldsymbol{X}}) \tag{3.32}$$

If the number of basic variable is n, we can obtain, by stationary condition of L, the following $n + 1$ equations with $n + 1$ unknowns

$$\frac{\partial L}{\partial \overline{\boldsymbol{X}}} = \frac{\overline{\boldsymbol{C}}^{-1} \overline{\boldsymbol{X}}}{(\overline{\boldsymbol{X}}^T \overline{\boldsymbol{C}}^{-1} \overline{\boldsymbol{X}})^{1/2}} + \lambda \overline{\boldsymbol{G}} = 0 \tag{3.33}$$

and

$$\frac{\partial L}{\partial \lambda} = \overline{G}(\overline{\mathbf{X}}) = 0 \tag{3.34}$$

The solution of the above equations $(\overline{\boldsymbol{X}}_L, \lambda_L)$ is the stationary point of L. Equation (33) and (34) are only necessary conditions for $(\overline{\boldsymbol{X}}_L, \lambda_L)$ to be a local minimum. If L is twice differentiable, and its Hessian is positive definite then $(\overline{\boldsymbol{X}}_L, \lambda_L)$ is a strict local minimum. Further if L is also convex $\overline{\boldsymbol{X}}_L$ will be the global minimum, namely the design point $\overline{X}^*$. From Eq. (33) we obtain

$$\overline{X}^* = -\lambda^* D^* \overline{C} \overline{G}_* \tag{3.35}$$

in which D^* is the global minimum of D. Substituting Eq. (35) into (30) yields

$$\lambda^* = \pm (\overline{G}_*^T \overline{C} \overline{G}_*)^{-1/2} \tag{3.36}$$

Pre-multiplying both sides of Eq. (35) by $\overline{G}_*^T$, we have

$$D^* = \frac{\left| \overline{\boldsymbol{G}}_*^T \overline{\boldsymbol{X}}^* \right|}{(\overline{\boldsymbol{G}}_*^T \overline{\boldsymbol{C}} \overline{\boldsymbol{G}}_*)^{1/2}} = \beta \tag{3.37}$$

Equations (35) through (37) result in

$$\overline{\boldsymbol{X}}^{*} = \pm \frac{\overline{\boldsymbol{C}}\,\overline{\boldsymbol{G}}_{*}}{(\overline{\boldsymbol{G}}_{*}^{T}\overline{\boldsymbol{C}}\,\overline{\boldsymbol{G}}_{*})^{1/2}}\beta \tag{3.38}$$

Since

$$\overline{\boldsymbol{G}}_{*} = \boldsymbol{s}\boldsymbol{G}_{*} \tag{3.39}$$

and

$$\overline{\boldsymbol{C}} = \boldsymbol{s}^{-1}\boldsymbol{C}(\boldsymbol{s}^{-1})^{T} = \boldsymbol{s}^{-1}\boldsymbol{C}\boldsymbol{s}^{-1} \tag{3.40}$$

we can express Eq. (37) and (38) in the original space as

$$D^{*} = \frac{\left|\boldsymbol{G}_{*}^{T}\left(\boldsymbol{m} - \boldsymbol{X}^{*}\right)\right|}{\left(\boldsymbol{G}_{*}^{T}\boldsymbol{C}\boldsymbol{G}_{*}\right)^{1/2}} = \beta \tag{3.41}$$

with

$$\boldsymbol{X}^{*} = \boldsymbol{m} + \boldsymbol{s}N^{*}\beta = \boldsymbol{m} \pm \frac{\boldsymbol{C}\boldsymbol{G}_{*}}{(\boldsymbol{G}_{*}^{T}\boldsymbol{C}\boldsymbol{G}_{*})^{1/2}}\beta \tag{3.42}$$

And

$$\boldsymbol{N}^{*} = \pm \frac{\boldsymbol{s}^{-1}\boldsymbol{C}\boldsymbol{G}_{*}}{(\boldsymbol{G}_{*}^{T}\boldsymbol{C}\boldsymbol{G}_{*})^{1/2}} \tag{3.43}$$

In Eq. (40) through (43) $\boldsymbol{C}$ is the covariance matrix of $\boldsymbol{X}$. As a matter of fact, $\boldsymbol{N}^{*}$ is the normal cosine vector at design point in U space. If its origin is in safe state the minus sign in Eq. (35) should be used. In structural reliability analysis $\boldsymbol{N}^{*}$ is often referred to as sensitivity factor. It is actually the derivative vector of reliability index with respect to U_i. So it can reflect the relative stochastic importance of basic variables. If N_i is much lower compared to others, then we might as well regard X_i as deterministic. No doubt, this will help us reduce the dimension in reliability analysis. It should be noted that Eq.16 and 17 are only a special case of Eq. (41) and (42). To summarize, the above procedure can be performed by the following iteration

Algorithm 3.1

1) Assume initial vector X_0^*.(usually $\boldsymbol{m}$ is taken);

2) Evaluate G_* and $\boldsymbol{N}^*$ at $\boldsymbol{X}_i^*$;

3) Form $\boldsymbol{X}_i^* = \boldsymbol{m} + \boldsymbol{s}\boldsymbol{N}^* \beta$;

4) Substitute the above X_i^* into $G(\boldsymbol{X})$ and solve for β;

5) Use the current β from step 4 to evaluate X_i^*;

6) Set $i = i + 1$;

7) Repeat step 2 through 6 until convergence is achieved.

This algorithm is easy to implement and useful to understand all the concepts we have thus far introduced. However, it is far from sufficient for engineering purpose.

3.4 Treatment of non-normal random variables

So far, we have established the basic first-order formulation from both geometric and optimization point of view with restrictions relaxed step by step. Now the only thing left to make the aforementioned procedure more general is the treatment of non-normal variables. This will involve the transformation of a general set of correlated non-normal variables in X space into an equivalent set of independent normal variables in U space. A closed form transformation for this purpose is the Rosenblatt transformation (Rosenblatt, 1952). This was first suggested by Hohenbichler and Rackwitz (1981).

Supposing $\boldsymbol{X}$ is a random vector with arbitrary joint c.d.f. $F(\boldsymbol{X})$, an independent standard normal vector $\boldsymbol{U}$ can be obtained from the following equations

$$\begin{aligned} U_1 &= \Phi^{-1}[F_1(X_1)] \\ U_2 &= \Phi^{-1}[F_2(X_2|X_1)] \\ &\vdots \\ U_n &= \Phi^{-1}[F_n(X_n|X_1,\cdots,X_{n-1})] \end{aligned} \tag{3.44}$$

or inversely

$$\begin{aligned} X_1 &= F_1^{-1}[\Phi(U_1)] \\ X_2 &= F_2^{-1}[\Phi(U_2)\big|X_1] \\ &\vdots \\ X_n &= F_n^{-1}[\Phi(U_n)\big|X_1, X_2, \cdots, X_{n-1}] \end{aligned} \tag{3.45}$$

They constitute the Rosenblatt transformation. The conditional c.d.f.'s in Eq. (44) and (45) can be expressed in terms of the joint p.d.f. as

$$F_i(X_i\big|X_1, X_2 \cdots X_{i-1}) = \frac{\int_{-\infty}^{X_i} f(X_1, X_2, \cdots, X_{i-1}, t)dt}{f(X_1, X_2, \cdots, X_{i-1})}, (i = 1,2,\ldots,n) \tag{3.46}$$

Unfortunately, the closed form of Eq. (46) can only be obtained in very special cases. Hohenbichler and Rackwitz (1981) suggested that they could be found by numerical differentiation. However such an approach is rather numerically unstable, and until now sees little application. In addition, it depends on the sequence of the random variables (Dolinski, 1983, Madsen, 1986).

To cast a new light on Rosenblatt transformation and make it more applicable, we might as well perform it by two steps: first transform a general random vector $\boldsymbol{X}$ into a dependent standard normal space of $\boldsymbol{Z}$, and then into the independent standard normal space of $\boldsymbol{U}$. The first step can be expressed as

$$Z_i = \Phi^{-1}[F_i(X_i)], (i = 1,2,\ldots n) \tag{3.47}$$

or

$$X_i = F_i^{-1}[\Phi(Z_i)], (i = 1,2,\ldots n) \tag{3.48}$$

and the second as, according to Eq. (22),

$$U = AZ \tag{3.49}$$

Thus the conditional c.d.f.'s are replaced by the marginal distribution which is relatively easy to obtain.

According to probabilistic theory the joint p.d.f. between any two random variables in X space can be expressed as

$$f_{ij}(X_i, X_j) = \phi_2(Z_i, Z_j, C_{zij}) \frac{f_i(X_i) f_j(X_j)}{\phi(Z_i)\phi(Z_j)} \tag{3.50}$$

where ϕ_2 is the joint p.d.f. of two standard normal variables, and C_{zij} is their coefficient of correlation. Similarly the joint p.d.f. of $\boldsymbol{X}$ has the form

$$f(\boldsymbol{X}) = \phi_n(\boldsymbol{Z}, \boldsymbol{C}_z) \frac{f_1(X_1) f_2(X_2) \ldots f_n(X_n)}{\phi(Z_1)\phi(Z_2)\ldots\phi(Z_n)} \tag{3.51}$$

where ϕ_n is the joint p.d.f. of n standard normal variables, and $\boldsymbol{C_z}$ is the corresponding correlation matrix. Equation (50) and (51) are traditionally called Nataf's model (Nataf,1962). So the coefficient of correlation in the original space can be determined by

$$\begin{aligned}
\rho_{Xij} &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(\frac{X_i - \mu_i}{\sigma_i}\right)\left(\frac{X_j - \mu_j}{\sigma_j}\right)\phi_2(Z_i, Z_j, C_{zij}) \frac{f_i(X_i) f_j(X_j)}{\phi(Z_i)\phi(Z_j)} dX_i dX_j \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(\frac{X_i - \mu_i}{\sigma_i}\right)\left(\frac{X_j - \mu_j}{\sigma_j}\right)\phi_2(Z_i, Z_j, C_{zij}) dZ_i dZ_j \\
&= \frac{E[X_i X_j] - \mu_i \mu_j}{\sigma_i \sigma_j}
\end{aligned} \tag{3.52}$$

Equation (52) can be solved by integration-iteration method. This is rather time consuming since it involves a two-fold integration. Some empirical formulae have been given by Der Kiureghian (1986) to reduce the computational cost at the expanse of accuracy. However the author of the present work believe it is not appropriate to solve Eq. (52) directly. Through some mathematical transformation we can reduce the problem to one much easier to solve. Let us express Eq. (48) by Hermite expansion as

$$X_i = \sum_{k=0}^{\infty} a_{ik} H_k(Z_i) \tag{3.53}$$

Here H_k(Z) is the k_{th} Hermite polynomial which satisfies the differentiation rule

$$\frac{d}{dZ} H_k(Z) = nH_{k-1}(Z) \tag{3.54}$$

and the orthogonal condition

$$\int_{-\infty}^{\infty} \phi(Z) H_m(Z) H_n(Z) dZ = n!\delta_{mn} \tag{3.55}$$

where δ_{mn} is delta function. So the coefficients a_{ik} in Eq. (53) can be determined by the inner product

$$a_{ik} = \frac{1}{k!} \int_{-\infty}^{\infty} F_i^{-1}[\Phi(Z_i)] H_k(Z_i) dZ_i \tag{3.56}$$

It follows that

$$E[X_i X_j] = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{im} a_{jn} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H_m(Z_i) H_n(Z_j) \phi_2(Z_i, Z_j, C_{zij}) dZ_i dZ_j \tag{3.57}$$

By Taylor expansion with respect to C_{zij} we obtain

$$E[X_i X_j] = \sum_{n=0}^{\infty} \frac{1}{n!} C_{zij}^n \left(\frac{d}{dC_{zij}}\right)^n E[X_i X_j] \Big|_{C_{zij}=0} \tag{3.58}$$

Since

$$\begin{aligned} &\frac{d}{dC_{zij}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H_m(Z_i) H_n(Z_j) \phi_2(Z_i, Z_j, C_{zij}) dZ_i dZ_j \\ &= mn \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H_{m-1}(Z_i) H_{n-}(Z_j) \phi_2(Z_i, Z_j, C_{zij}) dZ_i dZ_j \end{aligned} \tag{3.59}$$

Eq. (58) can be rewritten as

$$E[X_i X_j] = \sum_{n=0}^{\infty} n! a_{in} a_{jn} C_{zij}^n \tag{3.60}$$

Thus we reduce the original problem to a series of single integrals and the problem of solving a set of uncoupled nonlinear equation. A lot of standard methods are available to find the roots of Eq. (60), such as Newton iteration. The real root whose absolute value is less than or equal to 1 should be used as C_{zij}. Since the Hermite coefficients a_n decrease very fast with n, we can take $n = 10$ in practical calculation.

As we have noticed the transformation in Eq. (47) and (48) are non-linear, while Eq. (49) is a linear. This gives us a new way to view the problem: What will happen if we linearize it? To answer this question let us expand Eq. (47) at design point X_i^*:

$$Z_i = \Phi^{-1}[F_i(X_i^*)] + \frac{f_i(X_i^*)}{\phi\{\Phi^{-1}[F_i(X_i^*)]\}}(X_i' - X_i^*) \tag{3.61}$$

Since Z_i is a normal variable, X_i' must be normal too. It is termed equivalent normal variable. Considering that the statistics of Z_i should not be changed in linearization, and

$$\frac{f_i(X_i^*)}{\phi\{\Phi^{-1}[F_i(X_i^*)]\}} \geq 0 \tag{3.62}$$

we have

$$\sigma_{Z_i} = \frac{f_i(X_i^*)}{\phi\{\Phi^{-1}[F_i(X_i^*)]\}}\sigma_i' = 1 \tag{3.63}$$

$$\mu_{Z_i} = \Phi^{-1}[F_i(X_i^*)] + (\mu_i' - X_i^*)/\sigma_i' = 0 \tag{3.64}$$

Here μ_i' and σ_i' are the equivalent mean values and the equivalent standard deviations of.

$X_i^{'}$. With an eye to (3.41) and (3.42), we obtain

$$\sigma' = \frac{\phi\{\Phi^{-1}[F_i(X_i^*)]\}}{f_i(X_i^*)} \tag{3.65}$$

$$\mu' = X_i^* - \sigma_i' \Phi^{-1}[F_i(X_i^*)] \tag{3.66}$$

Finally, we can express the whole linearized transformation as

$$U = As'^{-1}(X' - m') \tag{3.67}$$

This means the iteration algorithm of FOSM can be performed (say the one in the last section) in the original space by substituting the equivalent normal variable $X_i^{'}$ for X_i at the current design point. In this way the original problem is reduced to a series of linear sub-problems. If the correlation matrix $\boldsymbol{C}_z$ is used, it will converge to the exact solution. As often as not the correlation matrices in X space and Z space are very close to each other. In such cases we can regard them as equal approximately to avoid solving Eq. (52). However as we will show in the examples, care must be taken.

3.5 General formulations of FORM by nonlinear programming

On the basis of the theoretical preparation in the preceding sections, it is now possible to give more applicable formulations for general purpose. This time we will present it in the context of nonlinear programming. The first attempt to introduce optimization method into FORM was made by Hasofer and Lind (1974). They used a simplified gradient algorithm. Later this method was improved to be globally convergent by Rackwitz and Fiessler (1978). The two methods are often quoted together as HL-RF method, which has become a standard method in FORM. Some modified versions of HL-RF method have been proposed (Abdo and Racwitz, 1990, Liu and Der Kiureghian, 1991). According to Liu and Der Kiureghian (1991), the frequently used optimization methods in structural reliability analysis are (1) gradient projection method; (2) penalty method; (3) augmented Lagrangian method; (4) modified HL-RF method; (5) sequential quadratic programming method (SQP). Their difference only resides in the determination of new search direction. Among them, the SQP has been widely accepted for its high robustness and efficiency. What is

more, it has been found that the HL-RF method and its modifications are but particular cases of SQP. Therefore it will be used in this section as well, but from a more general point of view than those in the aforementioned work.

SQP methods, also known as sequential or recursive, quadratic programming, employ Newton's method, or quasi-Newton methods to directly solve the Kuhn-Tucker condition of the original problem (Luenberger, 1973, Gill, 1981, and Bazaraa, et al. 1993). As a result, the accompanying sub-problem turns out to be the minimization of a quadratic approximation to the Lagrangian function optimized over a linear approximation to the constraints. Hence, this type of process is also known as a projected Lagrangian or Lagrangian-Newton approach. By its nature, this methods produceds both primal and dual (Lagrange multiplier) solutions. The most popular formats of SQP are those by Schittkowski (1980, 1981,1983, 1985).

To calculate the reliability index, consider the following problem in Z space

$$\begin{aligned} P1: \text{Minimize } & \frac{1}{2}\boldsymbol{Z}^T\boldsymbol{C}_z^{-1}\boldsymbol{Z} \\ \text{subject to } & G_z(\boldsymbol{Z}) = 0 \end{aligned} \tag{3.68}$$

Here we assume all functions are twice differentiable. The Kuhn-Tucker optimality conditions for problem $P1$ require a primal solution of $\boldsymbol{Z}^*$ and a Lagrange multiplier λ, to the effect that

$$\begin{aligned} &\nabla L(\boldsymbol{Z}) = \boldsymbol{C}_z^{-1}\boldsymbol{Z} + \lambda\nabla G_z(\boldsymbol{Z}) = 0 \\ &G_z(\boldsymbol{Z}) = 0 \end{aligned} \tag{3.69}$$

where $L(\mathbf{Z})$ is the Lagrangian function

$$L(\boldsymbol{Z},\lambda) = \frac{1}{2}\boldsymbol{Z}^T\boldsymbol{C}_z^{-1}\boldsymbol{Z} + \lambda G_z(\boldsymbol{Z}) \tag{3.70}$$

Equation (69) can be solved by Newton iteration

$$\begin{bmatrix} \boldsymbol{B} & \nabla G_z(\boldsymbol{Z}_k) \\ \nabla G_z(\boldsymbol{Z}_k) & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{d}_k \\ \lambda' \end{bmatrix} = \begin{bmatrix} -\boldsymbol{C}_z^{-1}\boldsymbol{Z}_{k+1} \\ -G_z(\boldsymbol{Z}_k) \end{bmatrix} \tag{3.71}$$

where $\boldsymbol{B}$ is the Hessian matrix of the Lagrangian function

$$\boldsymbol{B} = \nabla^2 L(\boldsymbol{Z}_k) = \boldsymbol{C}_z^{-1} + \lambda_k \nabla^2 G_z(\boldsymbol{Z}) \tag{3.72}$$

and

$$\boldsymbol{d}_k = \boldsymbol{Z}_{k+1} - \boldsymbol{Z}_k \tag{3.73}$$

Here $\boldsymbol{d}_k$ is called search direction. In fact Eq. (69) is also the Kuhn-Tucker condition of the following problem

$$\begin{aligned} P2: \text{Minimize } & L(\boldsymbol{Z}) + \nabla L(\boldsymbol{Z})^T \boldsymbol{d}_k + \frac{1}{2}\boldsymbol{d}_k{}^T \nabla^2 L(Z)\boldsymbol{d}_k \\ \text{subject to } & G_z(\boldsymbol{Z}) + \nabla G_z(\boldsymbol{Z})^T \boldsymbol{d}_k = 0 \end{aligned} \tag{3.74}$$

Observe that $P2$ represents a second-order Taylor series approximation for the Lagrangian function L and the first order linearization for the constraint. Compared to $P1$, $P2$ may be unbounded or infeasible in some cases.

The update of iteration point in optimization algorithms usually has the following form

$$\boldsymbol{Z}_{k+1} = \boldsymbol{Z}_k + \xi_k \boldsymbol{d}_k \tag{3.75}$$

$$\lambda_{k+1} = \lambda_k + \xi_k(\lambda' - \lambda_k) \tag{3.76}$$

where ξ_k (≥ 0) is the step length along the search direction. It is determined by minimizing a merit function in line search. This can be done through an exact method like quadratic fitting or inexact method such as Armijo's rule (Bazaraa, et al. 1993). Usually the augmented Lagrangian function is used as the merit function (Schittkowski, 1981a). In the present context it is

$$\Psi(\boldsymbol{Z},\lambda)=\frac{1}{2}\boldsymbol{Z}^T\boldsymbol{C}_z^{-1}\boldsymbol{Z}+\lambda G_z(\boldsymbol{Z})+\frac{1}{2}c_p G_z(\boldsymbol{Z})^2 \tag{3.77}$$

where c_p is the penalty factor used to ensure the convexity of ψ.

If Hessian matrix $\boldsymbol{B}$ is available Newton method has a convergence rate of order 2. However in practical structural reliability analysis it is often expensive or prohibitive to obtain even the first order derivatives of the limit state function, let alone the second order ones. Therefore $\boldsymbol{B}$ is usually updated by some approximate approaches. Broyden-Fletcher-Goldfarb-Shanno (BFGS) method is one of the most popular. It falls under the general class of quasi-Newton procedures. According to Schittkowski the BFGS scheme can be written as

$$\boldsymbol{B}_{k+1}=\boldsymbol{B}_k+\frac{\boldsymbol{q}\boldsymbol{q}^T}{\xi_k\boldsymbol{q}^T\boldsymbol{d}_k}-\frac{\boldsymbol{B}_k\boldsymbol{d}_k\boldsymbol{d}_k\boldsymbol{B}_k}{\boldsymbol{d}_k^T\boldsymbol{B}_k\boldsymbol{d}_k} \tag{3.78}$$

where

$$\boldsymbol{q}=\theta\boldsymbol{q}'+(1-\theta)\xi_k\boldsymbol{B}_k\boldsymbol{d}_k \tag{3.79}$$

with

$$\boldsymbol{q}'=\nabla L(\boldsymbol{Z}_{k+1},\lambda_{k+1})-\nabla L(\boldsymbol{Z}_k,\lambda_k) \tag{3.80}$$

and

$$\theta=\begin{cases}1 & \text{if } \boldsymbol{d}_k^T\boldsymbol{q}'\geq 0.2\xi_k\boldsymbol{d}_k^T\boldsymbol{B}_k\boldsymbol{d}_k\\ \dfrac{0.8\xi_k\boldsymbol{d}_k^T\boldsymbol{B}_k\boldsymbol{d}_k}{\xi_k\boldsymbol{d}_k^T\boldsymbol{B}_k\boldsymbol{d}_k-\boldsymbol{d}_k^T\boldsymbol{q}'} & \text{otherwise}\end{cases} \tag{3.81}$$

Schittkowski (1983) also suggested that the penalty factor c_p should be determined in an adaptive way to prevent ($\boldsymbol{Z}$, λ) from being infinity:

$$c_p^{(k)} = \max(c_p^{(k-1)}, \bar{c}^i) \tag{3.82}$$

Here $\bar{c}$ (> 1) is the tolerance of penalty factor, and i is the first positive integer, which satisfies

$$\frac{1}{\bar{c}^i} < \frac{1}{4} e_k \delta_k (1 - \frac{\delta_k}{4}), i = 1,2,\ldots \tag{3.83}$$

with

$$\delta_k = \min(\boldsymbol{d}_k^T \boldsymbol{B}_k \boldsymbol{d}_k / \| \boldsymbol{d}_k \|^2, \delta_{k-1}) \tag{3.84}$$

$$\delta_{-1} = 1 \tag{3.85}$$

$$e_k = \begin{cases} \| \boldsymbol{d}_k \|^2 / | \lambda_k - \lambda_{k-1} |^2 & \text{if } \lambda_k \neq \lambda_{k-1} \\ \bar{e} & \text{otherwise} \end{cases} \tag{3.86}$$

$$0 < \bar{e} < 1 \tag{3.87}$$

Now let us rewrite the Lagrangian function in $P2$ by linearizing the constraint

$$L(\boldsymbol{Z}, \lambda) = \frac{1}{2} \boldsymbol{Z}_k^T \boldsymbol{C}^{-1} \boldsymbol{Z}_k + \boldsymbol{C}^{-1} \boldsymbol{Z}_k \boldsymbol{d}_k + \frac{1}{2} \boldsymbol{d}_k^T \boldsymbol{C}^{-1} \boldsymbol{d}_k + \lambda (G_z(\boldsymbol{Z}_k) + \nabla G_z(\boldsymbol{Z}_k) \boldsymbol{d}_k) \tag{3.88}$$

The Kuhn-Tucker condition becomes:

$$\begin{aligned} &\nabla L(Z, \lambda) = \boldsymbol{C}_z^{-1} \boldsymbol{Z}_k + \boldsymbol{C}_z^{-1} \boldsymbol{d}_k + \lambda' \nabla G_z(\boldsymbol{Z}) = 0 \\ &G(\boldsymbol{Z}_k) + \nabla G(\boldsymbol{Z}_k) \boldsymbol{d}_k = 0 \end{aligned} \tag{3.89}$$

or in matrix form

$$\begin{bmatrix} \boldsymbol{C}_z^{-1} & \nabla G_z(\boldsymbol{Z}_k) \\ \nabla G_z(\boldsymbol{Z}_k) & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{d}_k \\ \lambda' \end{bmatrix} = \begin{bmatrix} -\boldsymbol{C}_z^{-1}\boldsymbol{Z}_k \\ -G_z(\boldsymbol{Z}_k) \end{bmatrix} \tag{3.90}$$

Apparently in this case the Hessian matrix $\boldsymbol{B}$ is replaced here by $\boldsymbol{C}_z^{-1}$, the inverse of correlation matrix in Z space. The solution of Eq. (90) can be easily found by inverting the coefficient matrix of the left-hand side

$$\begin{bmatrix} \boldsymbol{d}_k \\ \lambda' \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_z - \dfrac{\boldsymbol{C}_z\boldsymbol{G}_z\boldsymbol{G}_z^T\boldsymbol{C}_z^T}{\boldsymbol{G}_z^T\boldsymbol{C}_z\boldsymbol{G}_z} & \dfrac{\boldsymbol{G}_z^T\boldsymbol{C}_z^T}{\boldsymbol{G}_z^T\boldsymbol{C}_z\boldsymbol{G}_z} \\ \dfrac{\boldsymbol{G}_z^T\boldsymbol{C}_z^T}{\boldsymbol{G}_z^T\boldsymbol{C}_z\boldsymbol{G}_z} & \dfrac{-1}{\boldsymbol{G}_z^T\boldsymbol{C}_z\boldsymbol{G}_z} \end{bmatrix} \begin{bmatrix} -\boldsymbol{C}_z^{-1}\boldsymbol{Z}_k \\ -G_z(\boldsymbol{Z}_k) \end{bmatrix} \tag{3.91}$$

where for connivance we use $\boldsymbol{G}_z$ to denote $\nabla G_z(\boldsymbol{Z}_k)$. It follows that

$$\boldsymbol{d}_k = \frac{\boldsymbol{C}_z\boldsymbol{G}_z}{\boldsymbol{G}_z^T\boldsymbol{C}_z\boldsymbol{G}_z}(\boldsymbol{G}_z^T\boldsymbol{Z}_k - G_z(\boldsymbol{Z}_k)) - \boldsymbol{Z}_k \tag{3.92}$$

$$\lambda' = \frac{1}{\boldsymbol{G}_z^T\boldsymbol{C}_z\boldsymbol{G}_z}(G_z(\boldsymbol{Z}_k) - \boldsymbol{G}_z^T\boldsymbol{Z}_k) \tag{3.93}$$

It should be noticed that when the basic variables are independent, namely $\boldsymbol{C}_z = \boldsymbol{I}$, the solution above will reduce to the modified RF method by Abdo and Rackwitz (1990). Therefore Eq. (92) and (93) are more general. In particular when the step length in Eq. (75) and (76) is 1, we have

$$\boldsymbol{Z}_{k+1} = \frac{\boldsymbol{C}_z\boldsymbol{G}_z}{\boldsymbol{G}_z^T\boldsymbol{C}_z\boldsymbol{G}_z}(\boldsymbol{G}_z^T\boldsymbol{Z}_k - G_z(\boldsymbol{Z}_k)) \tag{3.94}$$

$$\lambda_{k+1} = \frac{1}{\boldsymbol{G}_z^T\boldsymbol{C}_z\boldsymbol{G}_z}(G_z(\boldsymbol{Z}_k) - \boldsymbol{G}_z^T\boldsymbol{Z}_k) \tag{3.95}$$

This is the generalized form of HL-RF method. A geometric illustration of this approach is shown in Fig.3.4 in *U* space.

Although Eq. (71) can be solved in the same way as in Eq. (90), it is highly recommended that Cholesky decomposition $\boldsymbol{LDL}^T$ should be used. This is because the update of $\boldsymbol{B}$ may be indefinite even if the true Hessian matrix is used. In that case we could simply inverse the sign of $\boldsymbol{D}$ to obtain a positive definite $\boldsymbol{B}$. Alternatively we might as well set $\boldsymbol{B}$ as $\boldsymbol{C}_z^{-1}$ in reliability analysis.

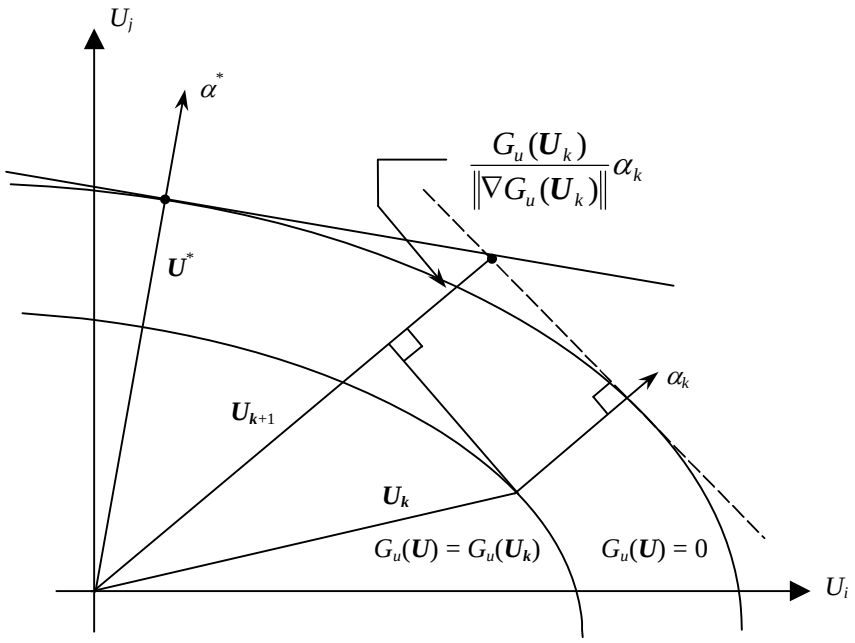


Figure 3.4 Illustration of HL-RF method

To conclude our discussion in this section, a general formulation of FORM is given below:

Algorithm 3.2

1) Assume initial vector $\boldsymbol{X}_0^*$, usually the mean vector $\boldsymbol{m}$ is used.

2) Calculate the corresponding failure point $\boldsymbol{Z}_0^*$ and reliability index β_0 in Z space.

3) Initialize parameters in SQP, c_{p0}, λ_0.

4) Determine the Jacobian matrix at current failure point.

$$\boldsymbol{J} = \frac{\partial(Z_1, Z_2, \ldots Z_n)}{\partial(X_1, X_2, \ldots X_n)}\Big|\boldsymbol{X}_k^* = \begin{bmatrix} \frac{\partial Z_1}{\partial X_1} & \frac{\partial Z_1}{\partial X_2} & \cdots & \frac{\partial Z_1}{\partial X_n} \\ \frac{\partial Z_2}{\partial X_1} & \frac{\partial Z_2}{\partial X_2} & \cdots & \frac{\partial Z_2}{\partial X_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial Z_n}{\partial X_1} & \frac{\partial Z_n}{\partial X_2} & \cdots & \frac{\partial Z_n}{\partial X_n} \end{bmatrix} \tag{3.96}$$

5) Evaluate the performance function and gradient vector at $\boldsymbol{Z}_k^*$

$$G_z(\boldsymbol{Z}_k^*) = G(\boldsymbol{X}_k^*) \tag{3.97}$$

$$\boldsymbol{G}_z^* = (\boldsymbol{J}^{-1})^T \boldsymbol{G}^* \tag{3.98}$$

6) Obtain a new failure point $\boldsymbol{Z}_{k+1}^*$ in Z space by one of the SQP methods.

7) Transform $\boldsymbol{Z}_{k+1}^*$ to the original space by Eq. (48) or its first-order approximation:

$$\boldsymbol{X}_{k+1}^* \approx \boldsymbol{X}_k^* + \boldsymbol{J}^{-1}(\boldsymbol{Z}_{k+1}^* - \boldsymbol{Z}_k^*) \tag{3.99}$$

8) Calculate safety index by

$$\beta_{k+1} = (\boldsymbol{Z}_{k+1}^{*T} \boldsymbol{C}_z^{-1} \boldsymbol{Z}_{k+1}^*)^{1/2} \tag{3.100}$$

9) Repeat step 4 through 8 until the following criteria are all satisfied:

$$|\beta_{k+1} - \beta_k| \le \varepsilon_1 \tag{3.101}$$

$$\| \boldsymbol{X}_{k+1}^* - \boldsymbol{X}_k^* \|^{1/2} \le \varepsilon_2 \tag{3.102}$$

$$|G(\boldsymbol{X}_{k+1})| \le \varepsilon_3 \tag{3.103}$$

where ε_1, ε_2, and ε_3 are tolerances defined by the user.

The Jacobian matrix of step 3 can be obtained through implicit differentiation

$$\frac{\partial Z_i}{\partial X_j} = \frac{\partial \Phi^{-1}[F_i(X_i \mid \ldots)]}{\partial X_j} = \frac{1}{\phi(Z_i)} \frac{\partial F_i(X_i \mid \ldots)}{\partial X_j} \tag{3.104}$$

If Rosenblatt transformation is used, the Jacobian matrix will be a lower triangular one (since $\partial U_i / \partial X_j = 0$ for $i < j$). Its inverse $\boldsymbol{J}^{-1}$ is easily obtained through back substitution. If stepwise transformation in section 3.4 is used $\boldsymbol{J}$ is a diagonal matrix.

Algorithm 3.2 is fairly versatile and suitable for a large range of engineering problems. By changing the Hessian matrix $\boldsymbol{B}$ in step 5, it can be shifted to either Quasi-Newton SQP method, or modified HL-RF method, or HL-RF method with robustness as well as computational cost in descending order. From the author's experience, Quasi-Newton SQP method is superior in small or medium sized problems with strong non-linearity in Z space. Modified HL-RF method is preferred in large size problem with certain non-linearity. If the non-linearity is not strong and the first order information is easy to obtain HL-RF method is the cheapest choice.

3.6 Sensitivity analysis

In structural design, it is important to know the stochastic importance of basic variables or the sensitivity of the failure probability or the reliability index to variations of parameters. By parameters here we mean either parameters of distribution or the deterministic parameters in the limit state function. Hohenbichler and Rackwitz (1986) gave a thorough discussion on this subject. As a matter of fact, in the frame we have thus far established it is quite straightforward to obtain all the above sensitivities. The sensitivity with respect to basic variables in the original space can be expressed as:

$$\frac{\partial \beta}{\partial \boldsymbol{X}} = \frac{\boldsymbol{J}}{2(\boldsymbol{Z}^T \boldsymbol{C}_z^{-1} \boldsymbol{Z})^{1/2}} [\boldsymbol{C}_z^{-1} + (\boldsymbol{C}_z^{-1})^T] \boldsymbol{Z} = \frac{\boldsymbol{J} \boldsymbol{C}_Z^{-1} \boldsymbol{Z}}{\beta} \tag{3.105}$$

In *Z* space we have

$$\frac{\partial \beta}{\partial \boldsymbol{Z}} = \frac{1}{2(\boldsymbol{Z}^T \boldsymbol{C}_z^{-1} \boldsymbol{Z})^{1/2}} [\boldsymbol{C}_z^{-1} + (\boldsymbol{C}_z^{-1})^T] \boldsymbol{Z} = \frac{\boldsymbol{C}_Z^{-1} \boldsymbol{Z}}{\beta} \tag{3.106}$$

And in *U* space we have

$$\frac{\partial \beta}{\partial \mathbf{U}} = \frac{(\mathbf{A}^{-1})^T}{2(\mathbf{Z}^T \mathbf{C}_z^{-1} \mathbf{Z})^{1/2}} [\mathbf{C}_z^{-1} + (\mathbf{C}_z^{-1})^T] \mathbf{Z} = \frac{\mathbf{A}\mathbf{Z}}{(\mathbf{Z}^T \mathbf{C}_z^{-1} \mathbf{Z})^{1/2}} = \frac{\mathbf{U}}{\beta} \tag{3.107}$$

Because the basic variables in *Z* space are dimensionless and the corresponding sensitivity is standardized, it can be used as a measure of the stochastic importance. Though the sensitivity in *U* space is also standardized, it is not recommended by the author to be used as an importance measure, since the transformation from *Z* to *U* is not unique. That is to say if the *i*th sensitivity factor is larger in *U* space, it does not necessarily mean the *i*th basic variable in *X* or *Z* space is more important.

The sensitivity with respect to distribution parameters can be written as

$$\frac{\partial \beta}{\partial P_{ij}} = \frac{\partial \beta}{\partial Z_i} \frac{\partial Z_i}{\partial P_{ij}} = \frac{\partial \beta}{\partial Z_i} \frac{1}{\phi(Z_i)} \frac{\partial F(X_i)}{\partial P_{ij}} \tag{3.108}$$

where P_{ij} is the *j*th distribution parameter of the *i*th random variable. Finally for the limit state function parameters we have

$$\frac{\partial \beta}{\partial P_i} = \left(\frac{\partial \beta}{\partial \boldsymbol{X}} \right)^T \frac{\partial \boldsymbol{X}}{\partial P_i} \tag{3.109}$$

where by implicit differentiation

$$\frac{\partial X_j}{\partial P_i} = -\frac{\partial G / \partial P_i}{\partial G / \partial X_j}, \; j = 1,2,\ldots N \tag{3.110}$$

In the same way we can obtain a simpler form in *U* space

$$\frac{\partial \beta}{\partial P_i} = N \frac{\partial G / \partial P_i}{\left\| \nabla G_u \right\|} \tag{3.111}$$

where N is the number of basic variables.

The concept of sensitivity has been extended to omission sensitivity, that is, the variation of β when a random number is replaced by a deterministic number (Madsen, 1988) and ignorance sensitivity when a deterministic number is replaced by a random variable (Der Kiureghian et.al., 1994; Maes, 1996).

3.7 Second order formulation

As is pointed out, the reliability description used in the first order formulation is a linear one by nature. Although it proves satisfactory when the limit state function is of pure or medium linearity, it is misleading when significant non-linearity is encountered. To limit the truncating error $|\Delta P_f|$ to a reasonable level, different second-order formulations have been proposed.

In second order method, the failure surface is replaced by a quadratic form in U space (Fiessler 1979). Since the characteristic function of the quadratic form is known the failure probability can be calculated by direct integration (Rice, 1980; Helstrom, 1983; Tvedt, 1990; Sakamoto, et al. 1997). Alternatively if the failure surface in rotated space is flat enough asymptotic integration can be employed (Breitung, 1984, 1989; Tvedt 1984, 1985; Cai and Elishakoff, 1994, Koyluoglu and Nielsen, 1994). The asymptotic failure probability appears as the first order failure probability multiplied by a coefficient in terms of the second order derivatives at the design point. It is easy to implement and more acceptable in structural design. All the above formulations will be discussed at length in this section.

3.7.1 Asymptotic approximation of Laplacian integral

As we have mentioned, the marrow of reliability analysis is to obtain the integral of joint p.d.f of basic variables in failure domain. In independent normal space, the integral can be reduced to Laplacian form as follows (see section 3.6.2)

$$I(\lambda) = \int_D p(\boldsymbol{x}) \exp[\lambda^2 h(\boldsymbol{x})] d\boldsymbol{x} \tag{3.112}$$

where $D \in R^n$, and λ is a large number.

According to large parameter theory, the greatest contribution to Laplace integrals is from the vicinity where the integrands have the maximum value. Assume the boundary of D is defined by

$$G = \{\boldsymbol{x} \big| g(\boldsymbol{x}) = 0\} \tag{3.113}$$

If $h(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are at least twice differentiable continuously and $h(\boldsymbol{x})$ only has one maximum point at $\boldsymbol{x}^*$ on G, then Eq. (112) can be expressed asymptotically as (Breitung, 1984)

$$I(\lambda) \to \frac{(2\pi)^{(n-1)/2}}{\lambda^{n+1}} \frac{p(\boldsymbol{x}^*) \exp[\lambda^2 h(\boldsymbol{x}^*)]}{\left|\boldsymbol{J}\right|^{1/2}}, \ (\lambda \to \infty) \tag{3.114}$$

with

$$\boldsymbol{J} = [\nabla h(\boldsymbol{x}^*)]^T \boldsymbol{B}(\boldsymbol{x}^*) \nabla h(\boldsymbol{x}^*) \tag{3.115}$$

$$\nabla h(\boldsymbol{x}^*) = \left(\frac{\partial h(\boldsymbol{x}^*)}{\partial x_1}, \frac{\partial h(\boldsymbol{x}^*)}{\partial x_1}, \cdots, \frac{\partial h(\boldsymbol{x}^*)}{\partial x_n} \right) \tag{3.116}$$

$$\boldsymbol{B}_{ij} = \operatorname{cof}\left(\frac{\partial^2 h(\boldsymbol{x}^*)}{\partial x_i \partial x_j} - K \frac{\partial^2 g(\boldsymbol{x}^*)}{\partial x_i \partial x_j} \right)_{n \times n} \tag{3.117}$$

$$K = \left\| \nabla h(\boldsymbol{x}^*) \right\| \Big/ \left\| \nabla g(\boldsymbol{x}^*) \right\| \tag{3.118}$$

In reliability analysis, G corresponds to failure surface and $\boldsymbol{x}^*$ to the design point. This confirms from another angle that failure probability is mainly subjected to the vicinity of design point.

3.7.2 SORM by asymptotic approximation

In U space, the second order Taylor expansion of failure function at the design point U^* is

$$G_S(\boldsymbol{U}) = \boldsymbol{G}_u^{*T}(\boldsymbol{U}-\boldsymbol{U}^*) + \frac{1}{2}(\boldsymbol{U}-\boldsymbol{U}^*)^T \boldsymbol{H}_u (\boldsymbol{U}-\boldsymbol{U}^*) \tag{3.119}$$

where $\boldsymbol{H}_u$ is the Hessian matrix in U space which is defined by

$$\boldsymbol{H}_u = \frac{\partial^2 \tilde{G}(\boldsymbol{U}^*)}{\partial U_i \partial U_j} \tag{3.120}$$

Divide Eq. (119) by $\left\|\nabla \tilde{G}(\boldsymbol{U}^*)\right\|$ we obtain its equivalent form

$$G_S(\boldsymbol{U}) = -\alpha^{*T}(\boldsymbol{U}-\boldsymbol{U}^*) + \frac{1}{2}(\boldsymbol{U}-\boldsymbol{U}^*)^T \boldsymbol{D}(\boldsymbol{U}-\boldsymbol{U}^*) \tag{3.121}$$

where

$$\alpha^* = -\frac{\nabla G_u(\boldsymbol{U}^*)}{\left\|\nabla G_u(\boldsymbol{U}^*)\right\|} \tag{3.122}$$

$$\boldsymbol{D} = \frac{\boldsymbol{H}_u}{\left\|\nabla G_u(\boldsymbol{U}^*)\right\|} \tag{3.123}$$

Defining orthogonal transformation

$$\boldsymbol{V} = \boldsymbol{R}\boldsymbol{U} \tag{3.124}$$

with the nth row in $\boldsymbol{R}$ equal to the unit gradient $\boldsymbol{\alpha}^*$, we can rewrite Eq. (121) in $\boldsymbol{V}$ space as

$$G_S(\boldsymbol{V}) = -(V_n - \beta) + \frac{1}{2}\begin{bmatrix} \hat{\boldsymbol{V}} \\ V_n - \beta \end{bmatrix}^T \boldsymbol{A} \begin{bmatrix} \hat{\boldsymbol{V}} \\ V_n - \beta \end{bmatrix} \tag{3.125}$$

where

$$\hat{\boldsymbol{V}} = (V_1, V_2, \ldots, V_{n-1})^T \tag{3.126}$$

and

$$\boldsymbol{A} = \boldsymbol{R}\boldsymbol{D}\boldsymbol{R}^T \tag{3.127}$$

Obviously, $\boldsymbol{V}$ is also a standard normal space with design point being

$$\boldsymbol{V}^* = (0,0,\ldots,\beta)^T \tag{3.128}$$

The failure probability in $\boldsymbol{V}$ space can be expressed as

$$P_f = (2\pi)^{-n/2} \int_{\tilde{G}(\boldsymbol{V})<0} \exp(-\frac{\|\boldsymbol{V}\|^2}{2}) d\boldsymbol{V} \tag{3.129}$$

Scaling $\boldsymbol{V}$ by

$$\overline{\boldsymbol{V}} = \frac{\boldsymbol{V}}{\beta} \tag{3.130}$$

produces

$$P_f = (2\pi)^{-n/2} \beta^n \int_{\tilde{G}(\beta\overline{\boldsymbol{V}})<0} \exp(-\frac{\beta^2 \|\overline{\boldsymbol{V}}\|^2}{2}) d\overline{\boldsymbol{V}} \tag{3.131}$$

Now we have an integral of Laplace type

$$\boldsymbol{I}(\beta) = \int_{\tilde{G}(\beta\overline{\boldsymbol{V}})<0} \exp(-\frac{\beta^2 \left\|\overline{\boldsymbol{V}}\right\|^2}{2}) d\overline{\boldsymbol{V}} \tag{3.132}$$

According to Eq. (114) through (118), its asymptotic form is

$$\boldsymbol{I}(\beta) \to (2\pi)^{(n-1)/2} \beta^{-(n+1)} \exp(\frac{-\beta^2}{2}) \left|\boldsymbol{J}\right|^{-1/2}, (\beta \to \infty) \tag{3.133}$$

with

$$\boldsymbol{J} = \frac{1}{\beta^2} \boldsymbol{V}^{*T} \boldsymbol{B}(\boldsymbol{V}^*) \boldsymbol{V}^* = B_{nn} \tag{3.134}$$

$$B_{nn} = \text{cof}(-\boldsymbol{I} - \beta \boldsymbol{A}) \tag{3.135}$$

It can be seen that by rotation transformation from U to V space, we reduce the calculation of n^2 determinants of (n-1) ×(n-1) matrix in Eq. (115) to only one. Substituting Eq. (133) into Eq. (131) yields

$$P_f \to (2\pi)^{-1/2} \beta^{-1} \exp(\frac{-\beta^2}{2}) \left|\boldsymbol{J}\right|^{-1/2}, \ (\beta \to \infty) \tag{3.136}$$

Considering Mill's ratio

$$\beta \to \frac{\phi(\beta)}{\Phi(-\beta)}, \ (\beta \to \infty) \tag{3.137}$$

we obtain

$$P_f \to \Phi(-\beta) \left|\boldsymbol{J}\right|^{-1/2} \tag{3.138}$$

This formula is first suggested by Breitung (1984). It takes the form of correction to the failure probability by FOSM, hence is widely used.

If we introduce matrix $\boldsymbol{A}'$ by

$$A'_{ij} = A_{ij}, (i, j = 1,2,\ldots,n-1) \tag{3.139}$$

The failure probability becomes

$$P_f = \Phi(-\beta)\prod_{i=1}^{n-1}(1+\beta\kappa_i) \tag{3.140}$$

where κ_i are the eigenvalues of matrix $\boldsymbol{A}'$. In fact the main curvature in V space along each axis is the eigenvalue of matrix $\boldsymbol{A}$ (Fiessler et al. 1979). As an approximation κ_i are often referred to as main or principle curvatures.

According to Eq. (125), for large β and small curvature problem, we have the following relationship

$$V_n \approx \beta + \frac{1}{2}\hat{\boldsymbol{V}}^T\boldsymbol{A}'\hat{\boldsymbol{V}} \tag{3.141}$$

Assume $G(\mathbf{0}) > 0$ in V space (otherwise, reliability measure will be obtained), the failure probability can be approximated by

$$P_f \approx \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\phi(V_1)\cdots\phi(V_{n-1})\int_{\beta+\frac{1}{2}\boldsymbol{VAV}}^{\infty}\phi(V_n)dV_n dV_1\cdots dV_{n-1} \tag{3.142}$$

Tvedt (1984) derived a three term approximation to Eq. (142) by a power series expansion in terms of $\hat{\boldsymbol{V}}^T\boldsymbol{A}'\hat{\boldsymbol{V}}$, ignoring terms of order higher than two. The result is

$$P_f \approx A_1 + A_2 + A_3 \tag{3.143}$$

with

$$A_1 = \Phi(-\beta)\prod_{i=1}^{n-1}(1+\beta\kappa_i)^{-1/2} \tag{3.144}$$

$$A_2 = [\beta\Phi(-\beta)-\phi(\beta)]\left\{\prod_{i=1}^{n-1}(1+\beta\kappa_i)^{-1/2} - \prod_{i=1}^{n-1}(1+(\beta+1)\kappa_i)^{-1/2}\right\} \tag{3.145}$$

$$A_3 = (\beta+1)[\beta\Phi(-\beta)-\phi(\beta)]\left\{\prod_{i=1}^{n-1}(1+\beta\kappa_i)^{-1.2} - \mathrm{RE}\left[\prod_{i=1}^{n-1}(1+(\beta+i)\kappa_i)^{-1/2}\right]\right\} \tag{3.146}$$

It can be seen that the first term of Tvedt's formula coincides with the Breitung's formula, so the former is more general. Good accuracy could be expected when the actual limit state in *V* space is close to a parabola.

To implement Eq. (140) and (143), it is required that matrix $\boldsymbol{I}+\beta\boldsymbol{A}'$ must be definite positive or

$$\beta\kappa_i > -1,\ (i=1,2,\ldots,n-1) \tag{3.147}$$

Otherwise singularity will happen. To avoid this, new approximation have been developped (Koyluoglu and Nielsen (1994); Cai and Elishakoff(1994)), based on McLaurin and Taylor expansion respectively. Cai's formula is expressed as

$$P_f = \Phi(-\beta)-\phi(\beta)(D_1+D_2+D_3) \tag{3.148}$$

with

$$D_1 = \sum_{i=1}^{n-1}\lambda_i \tag{3.149}$$

$$D_2 = -\frac{1}{2}\beta\left(3\sum_{i=1}^{n-1}\lambda_i^2 + \sum_{i\neq j}^{n-1}\lambda_i\lambda_j\right) \tag{3.150}$$

$$D_3 = \frac{1}{6}(\beta^2-1)\left(15\sum_{i=1}^{n-1}\lambda_i^3 + 9\sum_{i\neq j}^{n-1}\lambda_i^2\lambda_j + \sum_{i\neq j\neq m}^{n-1}\lambda_i\lambda_j\lambda_m\right) \tag{3.151}$$

and

$$\lambda_i = \kappa_i / 2,\ i = 1,2,\ldots,n-1 \tag{3.152}$$

Koyluoglu's formula is more complicated. According to the sign of main curvatures, it falls into three groups, each having one term, two term and three term approximation. The neglected contributions will approach to zero as $\beta \to \infty$. Since the two and three term approximations are too complicated for practical use, only one term approximation is given here. For positive curvatures, the approximation is

$$P_s = 1 - P_f = 1 - \Phi(\beta) \prod_{i=1}^{n-1} \frac{1}{\sqrt{1 + \kappa_i / c_{0,1}}} \tag{3.153}$$

where P_s is reliability measure and

$$c_{0,1} = \frac{\Phi(-\beta)}{\phi(\beta)} \tag{3.154}$$

For negative main curvatures, the approximation becomes

$$P_s = 1 - P_f = \Phi(\beta) \prod_{i=1}^{n-1} \frac{1}{\sqrt{1 - \kappa_i / c_{0,2}}} \tag{3.155}$$

where

$$c_{0,2} = \frac{\Phi(\beta)}{\phi(\beta)} \tag{3.156}$$

In the case of saddle points with positive and negative curvatures, assuming without loss of generality the first m main curvatures are positive and the rest are negative, the reliability measure becomes

$$P_s = 1 - P_f = \Phi(\beta) + \Phi(-\beta)\left[\prod_{i=m}^{n-1}\frac{1}{\sqrt{1-\kappa_i/d_{0,2}}}\left(1-\prod_{i=1}^{m-1}\frac{1}{\sqrt{1+\kappa_i/c_{0,1}}}\right)\right] - \Phi(\beta)\left[\prod_{i=1}^{m-1}\frac{1}{\sqrt{1+\kappa_i/d_{0,1}}}\left(1-\prod_{i=m}^{n-1}\frac{1}{\sqrt{1-\kappa_i/c_{0,2}}}\right)\right] \tag{3.157}$$

with

$$d_{0,1} = 2c_{0,1} \tag{3.158}$$

$$d_{02} = 2c_{0,2} \tag{3.159}$$

3.7.3 Distribution of quadratic form in normal space

Now let us give an insight into Eq. (125). If we define a new random variable M

$$M = V_n - \frac{1}{2}\begin{bmatrix}\hat{\boldsymbol{V}}\\ V_n-\beta\end{bmatrix}^T \boldsymbol{A}\begin{bmatrix}\hat{\boldsymbol{V}}\\ V_n-\beta\end{bmatrix} \tag{3.160}$$

the safety measure can be expressed as

$$P_s = P[G(\boldsymbol{V}) \ge 0] = P[M \le \beta] \tag{3.161}$$

To simplify Eq. (160), we can transform $\boldsymbol{V}$ into $\boldsymbol{Y}$ by

$$\boldsymbol{Y} = \boldsymbol{P}^T\boldsymbol{V} \tag{3.162}$$

whrere $\boldsymbol{P}$ is the eigenvector matrix of $\boldsymbol{A}$. This produces

$$M = -\frac{1}{2}A_{nn}\beta^2 + \boldsymbol{P}_n^T\boldsymbol{Y} + \beta\boldsymbol{P}_n^T\boldsymbol{\Lambda}\boldsymbol{Y} - \frac{1}{2}\boldsymbol{Y}^T\boldsymbol{\Lambda}\boldsymbol{Y} \tag{3.163}$$

where $\boldsymbol{P}_n$ is the nth row of $\boldsymbol{P}$ and $\boldsymbol{\Lambda}$ is the a diagonal matrix of eigenvalues of $\boldsymbol{A}$ (the actual

main curvatures). Obviously, M is a linear combination of a set of independent random variables, namely

$$M = v + \sum_{i=1}^{n} (\gamma_i Y_i - \frac{1}{2} \lambda_i Y_i^2) \tag{3.164}$$

with

$$v = -\frac{1}{2} A_{nn} \beta^2 \tag{3.165}$$

$$\boldsymbol{\gamma} = \boldsymbol{P}_n^T + \beta \boldsymbol{P}_n^T \Lambda \tag{3.166}$$

According to (Johnson and Kotz 1970), the characteristic function of M is

$$Q(\theta) = \exp(iv\theta) \prod_{j=1}^{n} \frac{\exp\left(-\frac{1}{2} \frac{\gamma_j^2 \theta^2}{1 + i\lambda_j \theta} \right)}{\sqrt{(1 + i\lambda_j \theta)}} \tag{3.167}$$

It constitutes a Fourier transform pair along with the p.d.f of M

$$f(M) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Q(\theta) \exp(-iM\theta) d\theta \tag{3.168}$$

$$Q(\theta) = \int_{-\infty}^{\infty} f(M) \exp(iM\theta) dM \tag{3.169}$$

This means the p.d.f of M can be calculated by FFT, and the c.d.f of M can be evaluated by trapezoidal integration (Sakamoto 1997)

$$F(M_k) = \sum_{j=1}^{K} \frac{f(M_j) + f(M_{j+1})}{2} \Delta M \tag{3.170}$$

If the integration step is properly chosen, safety measure can be obtained by Eq. (170) with very high precision.

In similar way we can define random variable from parabola in Eq. (141)

$$M_p = Y_n - \sum_{j=1}^{n-1} \frac{\kappa_j}{2} Y_j^2 \tag{3.171}$$

The corresponding characteristic function is

$$Q_p = \exp(-\frac{1}{2} i\theta^2) \prod_{j=1}^{n-1} \frac{1}{\sqrt{(1+i\kappa_j \theta)}} \tag{3.172}$$

Again FFT can be used to evaluate the safety measure. Another approach to calculate Eq. (168) is saddle integration (Rice 1980, Helstrom 1983, Tvedt 1990). However it is not as handy to implement.

3.7.4 SORM formulation by point-fitting

For the sake of mathematical perfection, we have assumed in the above discussion that the Hessian matrix of failure surface is known. In such case, the main curvatures can be obtained simply by eigenvalue analysis. This approach is usually referred to as curvature fitting. Obviously it is not suitable for those cases where the failure surface is not explicit, or the size of random variable is too large. To alleviate this problem, a point fitting scheme is proposed by Kiureghian et al.(1987), which has won great popularity.

In point fitting method, a simplified parabola in rotated standard space V is used instead of Eq. (141):

$$V_n \approx \beta + \frac{1}{2} \sum_{i=1}^{n-1} a_i V_i \tag{3.173}$$

where a_i are the main curvatures. The way to fit a_i is illustrated in Fig. 3.5. The basic idea is to make the probability content on the unsafe side of the intersecting parabola equal to

the sum of the probability contents of the two semi-parabolas. This can be expressed in the sense of Breitung's approximation as

$$\frac{1}{\sqrt{1+\beta a_i}} = \frac{1}{2}\left(\frac{1}{\sqrt{1+\beta a_{-i}}} + \frac{1}{\sqrt{1+\beta a_{+i}}}\right) \tag{3.174}$$

where

$$a_{\pm i} = 2(\eta_{\pm i} - \beta)/(k\beta)^2 \tag{3.175}$$

are curvatures of the two semi-parabolas. As a weighted version, a_i from Eq. (174) lies between a_{+i} and a_{-i} towards the one which corresponds to greater probability component. It can take into account automatically the effect of higher order terms in failure surface. In this respect it is superior to the curvature fitting method. The main curvature obtained is exact when the actual failure surface has a parabolic intersection.

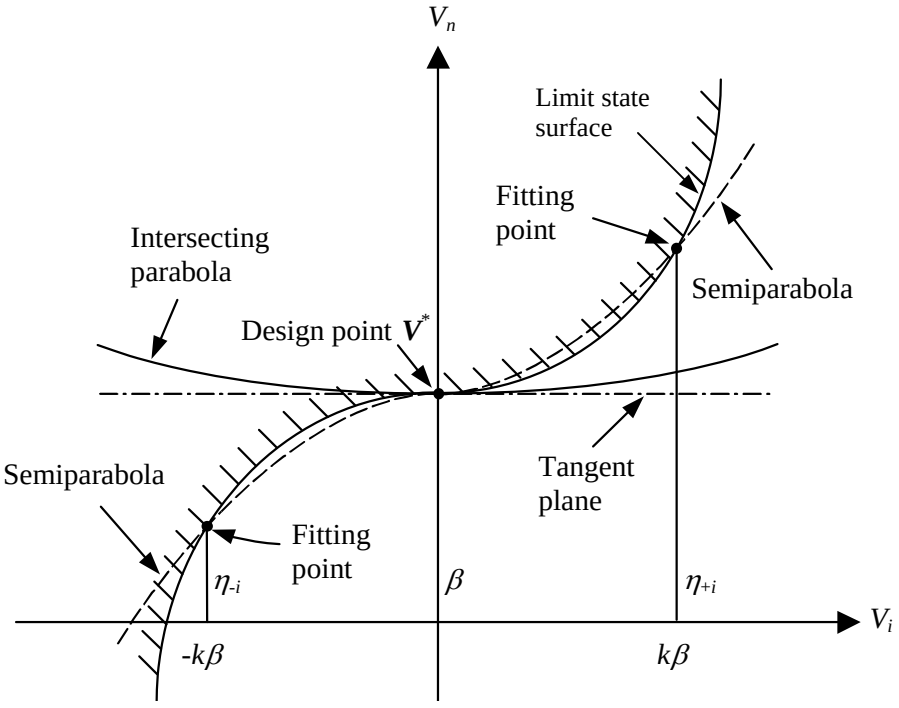


Fig.3.5 Fitting of paraboloid in rotated standard space

The coefficient k in Eq. (175) is used to control the location of fitting point in the neighbourhood of design point. Since the region of design point contributes more to the probability content, a too large k is not appropriate. The ratio of the probability density at a fitting point $(k\beta,\beta)$ to that at design point is $\exp[(k\beta)^2/2]$. If this ratio is set as 0.01, one obtains $k\beta = 3.03$. Considering that the reliability index in most engineering problems is around 3, k can be decided by the following rule due to Kiureghian et al. (1987):

$$k = \begin{cases} 1, & \text{for} \beta \le 3 \\ \dfrac{3}{\beta}, & \text{for} \beta > 3 \end{cases} \tag{3.176}$$

As to the coordinates $\eta_{\pm i}$ in Eq. (175), they can be calculated by any iteration method for solving nonlinear equation. By Newton method it can be written as

$$\eta_{k+1} = \eta_k - \frac{g(V_n^{(k)})}{g'(V_n^{(k)})} \tag{3.177}$$

which can ensure convergence at least of order 2. Alternatively, if the derivative information is difficult to obtain, the false-position method can be used:

$$\eta_{k+1} = \eta_k - g(V_n^{(k)}) \frac{\eta_k - \eta_{k-1}}{g(V_n^{(k)}) - g(V_n^{(k-1)})} \tag{3.178}$$

which is of order 1.618, the golden mean.

Compared with the complete form of parabola, the second order cross terms in Eq. (141) are neglected. The error thus introduced can be bounded through Breitung's formula:

$$\prod_{i=1}^{n-1}(1+\beta\kappa_i) \le \prod_{i=1}^{n-1}(1+\beta a_i) \le (1+\beta\bar{\kappa})^{n-1} \tag{3.179}$$

where $\bar{\kappa}$ is the average principal curvature:

$$\overline{\kappa} = \frac{\sum_{i=1}^{n-1} \kappa_i}{n-1} \tag{3.180}$$

Equation (179) reveals that the point-fitted paraboloid always underestimate the failure probability, and that the worst case is when the fitted main curvatures are equal to $\overline{\kappa}$.

In one of the further studies of Der Kiureghian et al. (1991), it is shown that the main curvatures can also be found by the final iteration information of the optimization methods converging towards the local minima. This rings a bell to us that in SQP the BFGS approximation of Hessian matrix will approach the actual one after certain iterations. However, in the author's point of view more studies are needed to apply the above results to SORM, especially when the number of random variable is large.

3.7.5 Generalized reliability index

The Hasofer-Lind reliability index β we have used hitherto is proposed under the background of first order analysis. It can not reflect the probability information of higher order. When used to compare the reliabilities of different structures, it often fails to give reasonable solution. To reserve the same concept while making it more adaptive, Ditlevsen (1979) proposed a generalised reliability index by

$$\beta_G = -\Phi^{-1}(1 - P_f) \tag{3.181}$$

This is, contrary to the Hasofer-Lind reliability index, an one-to-one mapping, which is also strictly monotonic namely

$$\beta_{G1} \le \beta_{G2} \Leftrightarrow P_{f1} \ge P_{f2} \tag{3.182}$$

Hereafter reliability index will still be denoted by β, which should be regarded as generalised one in the second order context. What is more, as β becomes infinitive, the variable and parameter sensitivity corresponding to general reliability index is asymptotically approaching that defined in section 3.6.

3.8 Examples

Based on the above derivation two modules (FORM, SORM) are written as a part of the structural system reliability analysis program GLAREL, which is developed by the author of the present work. Three examples are designed to test the new formulations.

3.8.1 Example 1

Consider the following limit state function

$$G(\boldsymbol{X}) = X_2 X_3 X_4 - \frac{X_5 X_3^2 X_4^2}{X_6 X_7} - X_1 = 0 \tag{3.183}$$

The mean vector of the basic variable is (0.01, 0.30, 360.0, 0.226×10^{-3}, 0.50, 0.12, 40.0), and the vector of c.o.v is (0.30, 0.05, 0.10, 0.05, 0.10, 0.05, 0.15). This example was originally used by Madsen (1986) to study the ultimate bending strength of a concrete beam with uncorrelated normal design variables. In order to make comparison, the example is extended: we still assume all variables have the same distribution, but four typical distributions are assigned in turn to the basic variables, namely normal distribution, lognormal distribution, extreme I distribution, and exponential distribution. HL-RF method is used in optimisation to obtain the reliability index. To begin with, first and second order reliability analyses are performed when the basic variables are independent. The results are listed in Table 3.1 through 3.6. Then a correlation matrix is assigned to the basic variables, namely

$$\boldsymbol{C} = \begin{bmatrix} 1.0 & 0.4 & 0.3 & 0.2 & 0.1 & 0.1 & 0.1 \\ 0.4 & 1.0 & 0.4 & 0.3 & 0.2 & 0.1 & 0.1 \\ 0.3 & 0.4 & 1.0 & 0.4 & 0.3 & 0.2 & 0.1 \\ 0.2 & 0.3 & 0.4 & 1.0 & 0.4 & 0.3 & 0.2 \\ 0.1 & 0.2 & 0.3 & 0.4 & 1.0 & 0.4 & 0.3 \\ 0.1 & 0.1 & 0.2 & 0.3 & 0.4 & 1.0 & 0.4 \\ 0.1 & 0.1 & 0.1 & 0.2 & 0.3 & 0.4 & 1.0 \end{bmatrix} \tag{3.184}$$

The corresponding results of reliability analysis are listed in Table 3.7 through 3.12.

It can be observed that though the limit state function has strong nonlinearity in the original space, it turns out to be rather linear in the vicinity of design point in *U* space, since the results of FORM do not differ much from those of SORM. In addition the distribution type plays an important role in reliability index. It decides the shape of failure surface in *U* space. Furthermore in the present case the introduction of correlation perturbs the final results dramatically towards safety. In second order analysis all the asymptotic formulae perform very well. When the reliability index is lager than 3 their precision is comparable to those from direct integration. Otherwise there will be a little deviation. The point-fitting scheme proves as effective as curvature-fitting approach but it is much faster than the latter. It deserves our notice that the second order reliability index from Taylor expansion is rather close to that from parabola approximation. This means that the effect of cross terms in Eq. (141) is negligible in rotated space *V*. Although these observations are only pertinent to the current example, they are enough to indicate that in real structural reliability analysis correlation between random variables should be considered where practically possible and the non-linearity of the failure surface at the design point should be checked by second order methods or Monte Carlo simulation.

Table 3.1 First order reliability analysis of Example 1

Distribution	Iteration Number	β	P_f	$G(X^*)$
Normal	10	3.413	0.3211×10^{-3}	-0.6939×10^{-17}
Lognormal	7	2.835	0.2291×10^{-2}	-0.3469×10^{-17}
Extreme I	12	2.713	3.337×10^{-2}	0.1388×10^{-16}
Exponential	18	2.443	0.7280×10^{-2}	0.3469×10^{-17}

Table 3.2 Design point of Example 1

Coordinates	Distribution			
	Normal	Lognormal	Extreme I	Exponential
X_1	0.01785	0.02069	0.02121	0.02138
X_2	0.2875	0.2928	0.2932	0.2930
X_3	292.8	328.5	336.2	339.8
X_4	0.2171×10^{-3}	0.2209×10^{-3}	0.2211×10^{-3}	0.2208×10^{-3}
X_5	0.5019	0.4987	0.4927	0.4853
X_6	0.1199	0.1198	0.1190	0.1181
X_7	39.66	39.35	38.85	38.05

Table 3.3 First order sensitivity analysis of Example 1

Directions	Distribution			
	Normal	Lognormal	Extreme I	Exponential
Z_1	0.7663	0.9254	0.9587	0.9805
Z_2	-0.2435	-0.1618	-0.1249	-0.8954×10^{-1}
Z_3	-0.5472	-0.3060	-0.2258	-0.1522
Z_4	-0.2316	-0.1533	-0.1187	-0.8533×10^{-1}
Z_5	0.1085×10^{-1}	0.8467×10^{-2}	0.7581×10^{-2}	0.6057×10^{-2}
Z_6	-0.5449×10^{-2}	-0.4242×10^{-2}	-0.3727×10^{-2}	-0.2934×10^{-2}
Z_7	-0.1647×10^{-1}	-0.1266×10^{-1}	-0.1133×10^{-1}	-0.8999×10^{-2}

Table 3.4 Second order reliability analysis of Example 1 by curvature fitting

Distribution	Formulae							
	Breitung		Tvedt		Cai		Koyluoglu	
	β	P_f	β	P_f	β	P_f	β	P_f
Normal	3.400	0.3373×10^{-3}	3.399	0.3385×10^{-3}	3.399	0.3384×10^{-3}	3.401	0.3362×10^{-3}
Lognormal	2.833	0.2305×10^{-2}	2.833	0.2306×10^{-2}	2.833	0.2306×10^{-2}	2.833	0.2306×10^{-2}
Extreme I	2.764	0.2849×10^{-2}	2.771	0.2795×10^{-2}	2.771	0.2794×10^{-2}	2.770	0.2803×10^{-2}
Exponential	2.553	0.5345×10^{-2}	2.569	0.5098×10^{-2}	2.571	0.5072×10^{-2}	2.565	0.5158×10^{-2}

Table 3.5 Second order reliability analysis of Example 1 by point fitting

Distribution	Formulae							
	Breitung		Tvedt		Cai		Koyluoglu	
	β	P_f	β	P_f	β	P_f	β	P_f
Normal	3.401	0.3361×10^{-3}	3.400	0.3372×10^{-3}	3.400	0.3372×10^{-3}	3.400	$.3366\times10^{-3}$
Lognormal	2.833	0.2305×10^{-2}	2.833	0.2307×10^{-2}	2.833	0.2307×10^{-2}	2.833	0.2307×10^{-2}
Extreme I	2.766	0.2840×10^{-2}	2.772	0.2785×10^{-2}	2.772	0.2784×10^{-2}	2.781	0.2701×10^{-2}
Exponential	2.537	0.5586×10^{-2}	2.551	0.5367×10^{-2}	2.551	0.5355×10^{-2}	2.585	0.4866×10^{-2}

Table 3.6 Second order reliability analysis of Example 1 by direct integration

Distribution	Curvature fitting				Point fitting	
	Parabola		Taylor expansion		Parabola	
	β	P_f	β	P_f	β	P_f
Normal	3.39876	0.338460×10^{-3}	3.39877	0.338450×10^{-3}	3.39980	0.337172×10^{-2}
Lognormal	2.83289	0.230648×10^{-2}	2.83157	0.231601×10^{-2}	2.83287	0.230659×10^{-2}
Extreme I	2.77088	0.279528×10^{-2}	2.76763	0.282328×10^{-2}	2.77207	0.278506×10^{-2}
Exponential	2.56869	0.510412×10^{-2}	2.56081	0.522137×10^{-2}	2.55096	0.537140×10^{-2}

Table 3.7 First order reliability analysis of Example 1 (dependent variables)

Distribution	Iteration Number	β	P_f	$G(X^*)$
Normal	14	3.905	0.4702×10^{-4}	-0.3469×10^{-17}
Lognormal	7	3.302	0.4798×10^{-3}	-0.3123×10^{-16}
Extreme I	19	3.111	0.9308×10^{-3}	0.3123×10^{-16}
Exponential	45	2.777	0.2746×10^{-2}	0.3469×10^{-16}

Table 3.8 Design point of Example 1 (dependent variables)

Coordinates	Distribution			
	Normal	Lognormal	Extreme I	Exponential
X_1	0.01647	0.02193	0.02267	0.02289
X_2	0.2866	0.3023	0.3027	0.3026
X_3	282.1	338.3	0.3470	0.3500
X_4	0.2080×10^{-3}	0.2199×10^{-3}	0.2216×10^{-3}	0.2221×10^{-3}
X_5	0.4517	0.4809	0.4828	0.4806
X_6	0.1660	0.1191	0.1188	0.1183
X_7	38.47	39.69	39.45	38.82

Table 3.9 First order sensitivity analysis of Example 1 (dependent variables)

Directions	Distribution			
	Normal	Lognormal	Extreme I	Exponential
Z_1	0.8949	1.078	1.096	1.113
Z_2	-0.2626	-0.1882	-0.1774	-0.1707
Z_3	-0.6140	-0.3568	-0.2857	-0.2415
Z_4	-0.2613	-0.1787	-0.1389	-0.1112
Z_5	0.1145×10^{-1}	0.9406×10^{-2}	0.7938×10^{-2}	0.6176×10^{-2}
Z_6	-0.5323×10^{-2}	-0.4712×10^{-2}	-0.4103×10^{-2}	-0.3375×10^{-2}
Z_7	-0.1613×10^{-1}	-0.1407×10^{-1}	-0.1289×10^{-1}	-0.1108×10^{-1}

Table 3.10 Second order reliability analysis of Example 1 by curvature fitting (dependent variables)

Distribution	Formulae							
	Breitung		Tvedt		Cai		Koyluoglu	
	β	P_f	β	P_f	β	P_f	β	P_f
Normal	3.915	0.4527×10^{-4}	3.915	0.4517×10^{-4}	3.915	0.4515×10^{-4}	3.915	0.4514×10^{-4}
Lognormal	3.300	0.4829×10^{-3}	3.300	0.4832×10^{-3}	3.300	0.4832×10^{-3}	3.300	0.4831×10^{-3}
Extreme I	3.182	0.7305×10^{-3}	3.189	0.7135×10^{-3}	3.190	0.7104×10^{-3}	3.188	0.7169×10^{-3}
Exponential	2.909	0.1822×10^{-2}	2.925	0.1722×10^{-2}	2.944	0.1620×10^{-2}	2.920	0.1748×10^{-2}

Table 3.11 Second order reliability analysis of Example 1 by point fitting (dependent variables)

Distribution	Formulae							
	Breitung		Tvedt		Cai		Koyluoglu	
	β	P_f	β	P_f	β	P_f	β	P_f
Normal	3.918	0.4461×10^{-4}	3.919	0.4447×10^{-4}	3.919	0.4447×10^{-4}	3.919	0.4443×10^{-4}
Lognormal	3.300	0.4830×10^{-3}	3.300	0.4832×10^{-3}	3.300	0.4832×10^{-3}	3.300	0.4832×10^{-3}
Extreme I	3.187	0.7184×10^{-3}	3.195	0.7001×10^{-3}	3.195	0.6991×10^{-3}	3.193	0.7031×10^{-3}
Exponential	2.915	0.1781×10^{-2}	2.932	0.1683×10^{-2}	2.936	0.1661×10^{-2}	2.928	0.1707×10^{-2}

Table 3.12 Second order reliability analysis of Example 1
by direct integration (dependent variables)

Distribution	Curvature fitting				Point fitting	
	Parabola		Taylor expansion		Parabola	
	β	P_f	β	P_f	β	P_f
Normal	3.915	0.4517×10^{-4}	3.915	0.4516×10^{-4}	3.919	0.4447×10^{-4}
Lognormal	3.300	0.4832×10^{-3}	3.076	0.1050×10^{-2}	3.300	0.4832×10^{-3}
Extreme I	3.189	0.7137×10^{-3}	3.125	0.8885×10^{-3}	3.194	0.7002×10^{-3}
Exponential	2.924	0.1725×10^{-2}	2.895	0.1896×10^{-2}	2.932	0.1685×10^{-2}

3.8.2 Example 2

Consider the limit state function generated by response-surface fitting

$$\begin{aligned} G(\boldsymbol{X}) = {} & 1.1 - 0.00534X_1 - 0.0705X_2 - 0.226X_3 + 0.998X_4 \\ & + 0.00117X_1^2 + 0.00157X_2^2 + 0.0333X_3^2 - 1.339X_4^2 \\ & - 0.00115X_1X_2 - 0.0149X_1X_3 + 0.0717X_1X_4 \\ & + 0.01350X_2X_3 - 0.0611X_2X_4 - 0.5580X_3X_4 \end{aligned} \tag{3.185}$$

The mean vector of basic variables is (10, 25, 0.8, 0.0625), and the vector of c.o.v is (0.5, 0.2, 0.25, 1.0). In this case, the failure surface is highly non-linear in *U* space. As a result HL-RF method failed to converge because of the oscillation in iteration. So the modified HL-RF method with Armijo's line search is used. The reliability analysis is performed in the same way as in Example 1. For convenience, the first four rows and columns of $\boldsymbol{C}$ in Eq. (184) is chosen as the current correlation matrix. The results are listed in Table 3.13 through 3.24.

This time we can observe significant difference between the results of FORM and SORM, which is a consequence of the strong non-linearity. For the same reason a small *k* in Eq. (175) has to be used to calculate the *n*th coordinate of fitting point. Otherwise, it is very difficult for the Newton method or Falsehood method to converge. Here *k* is chosen between 0.3-0.5. As is shown by the calculation, it leads to loss of probability information. The introduction of correlation makes the reliability index decrease a lot. This is contrary

to its effect in Example 1. In addition it has more bearing on the sensitivity of basic variables. In this example, both the first order reliability index and the radii of main curvature are at a low level. As a result, Cai's asymptotic formula in SORM behaves poorly. Fortunately, the other formulae still can reflect the probability information to certain precision.

Table 3.13 First order reliability analysis of Example 2

Distribution	Iteration Number	β	P_f	$G(X^*)$
Normal	24	0.9562	0.1695	0.3064×10^{-12}
Lognormal	24	1.159	0.1233	0.1586×10^{-14}
Extreme I	28	1.146	0.1259	0.9587×10^{-13}
Exponential	33	1.259	0.1040	0.6073×10^{-12}

Table 3.14 Design point of Example 2

Coordinates	Distribution			
	Normal	Lognormal	Extreme I	Exponential
X_1	14.22	15.08	14.98	15.45
X_2	25.31	24.94	24.93	24.60
X_3	0.8677	0.8424	0.8436	0.8181
X_4	0.08063	0.04515	0.05769	0.04156

Table 3.15 First order sensitivity analysis of Example 2

Directions	Distribution			
	Normal	Lognormal	Extreme I	Exponential
Z_1	0.8823	0.9547	0.9237	0.9191
Z_2	0.06399	0.07529	0.1419	0.2040
Z_3	0.3540	0.2872	0.3460	0.3360
Z_4	0.3034	0.02220	0.08351	-0.02853

Table 3.16 Second order reliability analysis of Example 2 by curvature fitting

Distribution	Formulae							
	Breitung		Tvedt		Cai		Koyluoglu	
	β	P_f	β	P_f	β	P_f	β	P_f
Normal	1.213	0.1125	1.291	0.09830	0.4313	0.3331	1.272	0.1016
Lognormal	1.442	0.07452	1.552	0.06034	1.288	0.09880	1.491	0.06792
Extreme I	1.361	0.08670	1.401	0.08062	1.151	0.1249	1.409	0.07939
Exponential	1.527	0.06339	1.605	0.05427	1.869	0.03083	1.562	0.05918

Table 3.17 Second order reliability analysis of Example 2 by point fitting

Distribution	Formulae							
	Breitung		Tvedt		Cai		Koyluoglu	
	β	P_f	β	P_f	β	P_f	β	P_f
Normal	1.285	0.09937	1.416	0.07829	0.2615	0.3968	1.350	0.08552
Lognormal	1.516	0.06471	1.664	0.04809	1.305	0.09593	1.580	0.05711
Extreme I	1.462	0.07193	1.567	0.05854	1.186	0.1178	1.512	0.06524
Exponential	1.643	0.05016	1.785	0.03710	2.149	0.01582	1.708	0.04384

Table 3.18 Second order reliability analysis of Example 2 by direct integration

Distribution	Curvature fitting				Point fitting	
	Parabola		Taylor expansion		Parabola	
	β	P_f	β	P_f	β	P_f
Normal	1.296	0.09755	1.378	0.08412	1.391	0.08199
Lognormal	1.525	0.06358	1.624	0.05222	1.625	0.05204
Extreme I	1.416	0.07840	1.490	0.06811	1.545	0.06113
Exponential	1.586	0.05639	1.667	0.04777	1.750	0.04007

Table 3.19 First order reliability analysis of Example 2 (dependent variables)

Distribution	Iteration Number	β	P_f	$G(X^*)$
Normal	9	0.8489	0.1980	-0.4454×10^{-12}
Lognormal	10	1.088	0.1384	-2.528×10^{-12}
Extreme I	14	0.9999	0.1587	-0.2347×10^{-12}
Exponential	17	1.088	0.1382	-0.3617×10^{-13}

Table 3.20 Design point of Example 2 (dependent variables)

Coordinates	Distribution			
	Normal	Lognormal	Extreme I	Exponential
X_1	13.75	14.62	14.08	14.40
X_2	26.33	26.23	26.09	25.90
X_3	0.8910	0.8840	0.8772	0.8656
X_4	0.09272	0.06391	0.08424	0.07541

Table 3.21 First order sensitivity analysis of Example 2 (dependent variables)

Directions	Distribution			
	Normal	Lognormal	Extreme I	Exponential
Z_1	0.8295	0.9466	0.8466	0.8563
Z_2	-0.2234	-0.2219	-0.1410	-0.09556
Z_3	0.2238	0.2009	0.2267	0.2316
Z_4	0.3812	0.1629	0.2912	0.2135

Table 3.22 Second order reliability analysis of Example 2 by curvature fitting (dependent variables)

Distribution	Formulae							
	Breitung		Tvedt		Cai		Koyluoglu	
	β	P_f	β	P_f	β	P_f	β	P_f
Normal	1.042	0.1487	1.161	0.1228	0.8451	0.1991	1.102	0.1353
Lognormal	1.220	0.1112	1.264	0.1031	1.153	0.1245	1.259	0.1040
Extreme I	1.117	0.1319	1.144	0.1263	0.9988	0.1589	1.161	0.1228
Exponential	1.180	0.1190	1.151	0.1248	1.085	0.1390	1.229	0.1095

Table 3.23 Second order reliability analysis of Example 2 by point fitting (dependent variables)

Distribution	Formulae							
	Breitung		Tvedt		Cai		Koyluoglu	
	β	P_f	β	P_f	β	P_f	β	P_f
Normal	1.088	0.1382	1.264	0.1030	1.086	0.1388	2.481	0.006558
Lognormal	1.248	0.1059	1.315	0.09420	1.191	0.1168	2.126	0.01675
Extreme I	1.180	0.1190	1.274	0.1013	1.174	0.1202	1.862	0.03128
Exponential	1.251	0.1054	1.316	0.09402	1.236	0.1082	1.904	0.02843

Table 3.24 Second order reliability analysis of Example 2
by direct integration (dependent variables)

Distribution	Curvature fitting				Point fitting	
	Parabola		Taylor expansion		Parabola	
	β	P_f	β	P_f	β	P_f
Normal	1.132	0.1288	1.195	0.1161	1.227	0.1100
Lognormal	1.263	0.1033	1.319	0.09362	1.301	0.09662
Extreme I	1.160	0.1230	1.200	0.1151	1.257	0.1043
Exponential	1.215	0.1123	1.268	0.1024	1.313	0.09466

3.8.3 Example 3

Consider a linear polynomial limit state with sinusoidal noise

$$G(\boldsymbol{X}) = X_1 + 2X_2 + 2X_3 + X_4 - 5X_5 - 5X_6 + 0.001\sum_{i=1}^{6}\sin(100X_i) \qquad (3.186)$$

This example was first used by Der Kiureghian (1987) as a variation of the example used by Tvdet (1984), representing failure in one plastic collapse mechanism of a one bay frame. Here we further extend it by using different distribution and assumed correlation. The reliability analysis is carried out in the same way as in the previous two examples. The first 6 order sub-matrix of $\boldsymbol{C}$ in Eq. (184) is used as the current correlation matrix. The existence of noise makes the problem so worse that it severely changes the nonlinearity of the original problem. Consequently, HL-RF method fails totally in this example. And it takes the modified HL-RF method hundreds of iterations to converge, when the basic variables are independent. Otherwise it fails as well. Only SQP method survives through such a tricky test with perfect performance. However, this is not the end of story. It is found that there is more than one minimum point in *U* space. Since all the minima tend to give more or less the same reliability index and they are close to each other, it is reasonable to believe that they are approximately distributed on a small hyper-sphere. Listed in Table 3.25 through 3.36 are only the results concerning one minimum point. The reliability indexes of the original problem (without noise) are listed in Table 3.37 and Table 3.38. The variation of main curvatures is given by Table 3.39 and 3.40.

It can be noticed that the main curvatures of the original problem are drastically distorted by the noise. As a result, the reliability obtained by curvature fitting deviates a lot from the target problem. In contrast, the point fitting approach is not sensitive to the noise and gives very good accuracy. Furthermore, though the first order reliability index is at high level, the error of asymptotic integration is significant because of the over enlarged main curvatures. The same reason accounts for the singularity that happens to Cai's formula. According to the author's experience in some cases singularity may happen to Koyluoglu's formula as well. These show that they are not ideal replacements for Breitung's and Tvedt's formulae.

Table 3.25 First order reliability analysis of Example 3

Distribution	Iteration Number	β	P_f	$G(X^*)$
Normal	25	2.615	0.4467×10^{-2}	0.1162×10^{-12}
Lognormal	23	2.348	0.9432×10^{-2}	0.8718×10^{-12}
Extreme I	35	2.287	0.1111×10^{-1}	0.4788×10^{-12}
Exponential	32	2.156	0.1554×10^{-1}	0.4310×10^{-12}

Table 3.26 Design point of Example 3

Coordinates	Distribution			
	Normal	Lognormal	Extreme I	Exponential
X_1	116.4	117.2	116.7	115.4
X_2	112.5	115.2	115.2	114.7
X_3	112.5	115.2	115.2	114.7
X_4	116.4	117.2	116.6	115.3
X_5	78.53	83.38	84.93	87.39
X_6	58.05	55.66	53.89	50.51

Table 3.27 First order sensitivity analysis of Example 3

Directions	Distribution			
	Normal	Lognormal	Extreme I	Exponential
Z_1	-0.1144	-0.08069	-0.05723	-0.04887
Z_2	-0.2384	-0.1525	-0.1197	-0.8380
Z_3	-0.2384	-0.1525	-0.1197	-0.8380
Z_4	-0.1144	-0.08069	-0.5963	-0.4887
Z_5	0.7274	0.8044	0.8362	0.8696
Z_6	0.5753	0.5418	0.5150	0.4744

Table 3.28 Second order reliability analysis of Example 3 by curvature fitting

Distribution	Formulae							
	Breitung		Tvedt		Cai		Koyluoglu	
	β	P_f	β	P_f	β	P_f	β	P_f
Normal	4.856	0.5998×10^{-6}	5.086	0.1825×10^{-6}	Singular		4.910	0.4556×10^{-6}
Lognormal	4.400	0.5397×10^{-5}	4.722	0.1167×10^{-5}	Singular		4.469	0.3925×10^{-5}
Extreme I	4.259	0.1025×10^{-4}	4.595	0.2160×10^{-5}	Singular		4.331	0.7410×10^{-5}
Exponential	3.884	0.5132×10^{-4}	4.275	0.9572×10^{-5}	Singular		3.967	0.3644×10^{-4}

Table 3.29 Second order reliability analysis of Example 3 by point fitting

Distribution	Formulae							
	Breitung		Tvedt		Cai		Koyluoglu	
	β	P_f	β	P_f	β	P_f	β	P_f
Normal	2.614	0.4469×10^{-2}	2.614	0.4469×10^{-2}	2.614	0.4469×10^{-2}	2.614	0.4469×10^{-2}
Lognormal	2.262	0.1185×10^{-1}	2.252	0.1216×10^{-1}	2.255	0.1206×10^{-1}	2.243	0.1244×10^{-1}
Extreme I	2.208	0.1361×10^{-1}	2.199	0.1395×10^{-1}	2.206	0.1371×10^{-1}	2.180	0.1461×10^{-1}
Exponential	2.074	0.1904×10^{-1}	2.054	0.1996×10^{-1}	2.078	0.1886×10^{-1}	2.026	0.2140×10^{-1}

Table 3.30 Second order reliability analysis of Example 3 by direct integration

Distribution	Curvature fitting				Point fitting	
	Parabola		Taylor expansion		Parabola	
	β	P_f	β	P_f	β	P_f
Normal	0.8897	0.1868	1.075	0.1413	2.614	0.4469×10^{-2}
Lognormal	1.506	0.6600×10^{-1}	0.5528	0.2902	2.254	0.1208×10^{-1}
Extreme I	1.290	0.9844×10^{-1}	3.488	0.3636	2.208	0.1361×10^{-1}
Exponential	1.362	0.8656×10^{-1}	0.9547	0.1699	2.077	0.1891×10^{-1}

Table 3.31 First order reliability analysis of Example 3 (dependent variables)

Distribution	Iteration Number	β	P_f	$G(X^*)$
Normal	22	2.523	0.5814×10^{-2}	0.3403×10^{-12}
Lognormal	24	2.194	0.1411×10^{-1}	0.7842×10^{-13}
Extreme I	30	2.116	0.1718×10^{-1}	-0.7693×10^{-13}
Exponential	35	1.959	0.2507×10^{-1}	0.7510×10^{-12}

Table 3.32 Design point of Example 3 (dependent variables)

Coordinates	Distribution			
	Normal	Lognormal	Extreme I	Exponential
X_1	115.0	117.9	117.4	116.5
X_2	114.1	118.0	117.7	117.2
X_3	117.7	121.3	120.6	120.0
X_4	124.7	126.5	125.5	124.6
X_5	78.59	81.73	81.60	81.23
X_6	62.07	62.88	62.32	61.88

Table 3.33 First order sensitivity analysis of Example 3 (dependent variables)

Directions	Distribution			
	Normal	Lognormal	Extreme I	Exponential
Z_1	-0.1126	-0.07040	-0.05705	-0.04319
Z_2	-0.2209	-0.1460	-0.1206	-0.09160
Z_3	-0.2302	-0.1483	-0.1333	-0.1120
Z_4	-0.1124	-0.07855	-0.07182	-0.07105
Z_5	0.7002	0.7180	0.7212	0.7148
Z_6	0.5618	0.5509	0.5457	0.5438

Table 3.34 Second order reliability analysis of Example 3 by curvature fitting (dependent variables)

Distribution	Formulae							
	Breitung		Tvedt		Cai		Koyluoglu	
	β	P_f	β	P_f	β	P_f	β	P_f
Normal	4.688	0.1379×10^{-5}	4.943	0.3843×10^{-6}		Singular	4.747	0.1034×10^{-5}
Lognormal	4.223	0.1208×10^{-4}	4.614	0.1978×10^{-5}		Singular	4.300	0.8538×10^{-5}
Extreme I	4.160	0.1590×10^{-4}	4.611	0.1999×10^{-5}		Singular	4.243	0.1102×10^{-4}
Exponential	3.969	0.3602×10^{-4}	4.590	0.2218×10^{-5}		Singular	4.064	0.2408×10^{-4}

Table 3.35 Second order reliability analysis of Example 3 by point fitting (dependent variables)

Distribution	Formulae							
	Breitung		Tvedt		Cai		Koyluoglu	
	β	P_f	β	P_f	β	P_f	β	P_f
Normal	2.523	0.5813×10^{-2}	2.523	0.5813×10^{-2}	2.523	0.5813×10^{-2}	2.523	0.5813×10^{-2}
Lognormal	2.155	0.1560×10^{-1}	2.149	0.1584×10^{-1}	2.149	0.1583×10^{-1}	2.152	0.1568×10^{-1}
Extreme I	2.108	0.1753×10^{-1}	2.106	0.1759×10^{-1}	2.106	0.1759×10^{-1}	2.107	0.1758×10^{-1}
Exponential	1.977	0.2402×10^{-1}	1.981	0.2382×10^{-1}	1.981	0.2382×10^{-1}	1.980	0.2384×10^{-1}

Table 3.36 Second order reliability analysis of Example 3 by direct integration (dependent variables)

Distribution	Curvature fitting				Point fitting	
	Parabola		Taylor expansion		Parabola	
	β	P_f	β	P_f	β	P_f
Normal	0.7412	0.2293	0.8366	0.2014	2.523	0.5813×10^{-2}
Lognormal	1.494	0.6754×10^{-1}	0.3858	0.3498	2.149	0.1583×10^{-1}
Extreme I	1.543	0.6141×10^{-1}	0.3298	0.3708	2.106	0.1759×10^{-1}
Exponential	1.047	0.1475	0.8447	0.1991	1.983	0.2370×10^{-1}

Table 3.37 Second order reliability analysis of the original problem in Example 3 by direct integration

Distribution	Curvature fitting				Point fitting	
	Parabola		Taylor expansion		Parabola	
	β	P_f	β	P_f	β	P_f
Normal	2.614	0.4468×10^{-2}	2.614	0.4468×10^{-2}	2.614	0.4468×10^{-2}
Lognormal	2.250	0.1222×10^{-1}	2.241	0.1253×10^{-2}	2.254	0.1208×10^{-1}
Extreme I	2.199	0.1395×10^{-1}	2.176	0.1478×10^{-1}	2.245	0.1373×10^{-1}
Exponential	2.076	0.1895×10^{-1}	2.032	0.2107×10^{-1}	2.076	0.1892×10^{-1}

Table 3.38 Second order reliability analysis of the original problem in Example 3 by direct integration (dependent variables)

Distribution	Curvature fitting				Point fitting	
	Parabola		Taylor expansion		Parabola	
	β	P_f	β	P_f	β	P_f
Normal	2.523	0.5813×10^{-2}	2.523	0.5813×10^{-2}	2.523	0.5813×10^{-2}
Lognormal	2.143	0.1606×10^{-1}	2.141	0.1615×10^{-1}	2.148	0.1584×10^{-1}
Extreme I	2.095	0.1811×10^{-1}	2.091	0.1828×10^{-1}	2.106	0.1762×10^{-1}
Exponential	1.969	0.2449×10^{-1}	1.959	0.2506×10^{-1}	1.982	0.2372×10^{-1}

Table 3.39 Effect of noise on main curvatures in Example 3

Main Curvatures	Distribution							
	Normal		Lognormal		Extreme I		Exponential	
	Without noise	With noise	Without noise	With noise	Without noise	With noise	Without noise	With noise
κ_1	0	7.045	-0.2375	-10.17	-0.2936	2.420	-0.3374	3.062
κ_2	0	13.11	-0.1574	8.674	-0.1916	5.503	-0.2471	3.062
κ_3	0	13.11	0.007689	8.674	0.02183	5.909	0.03780	3.124
κ_4	0	13.69	0.007689	8.738	0.02183	5.909	0.03780	3.124
κ_5	0	13.69	0.01511	8.738	0.04099	15.21	0.07318	6.911
κ_6	0	20.91	0.01511	15.02	0.04099	45.07	0.07318	25.01

Table 3.40 Effect of noise on main curvatures in Example 3 (dependent variables)

Main Curvatures	Distribution							
	Normal		Lognormal		Extreme I		Exponential	
	Without noise	With noise	Without noise	With noise	Without noise	With noise	Without noise	With noise
κ_1	0	6.790	-0.2624	2.689	-0.3156	3.209	-0.3670	2.171
κ_2	0	7.530	-0.1065	4.315	-0.1294	3.808	-0.1424	3.394
κ_3	0	8.532	0.005071	6.071	0.01477	5.744	0.02402	5.227
κ_4	0	9.510	0.005186	11.09	0.01836	10.85	0.02614	10.37
κ_5	0	16.32	0.008614	13.84	0.02636	15.08	0.04173	16.28
κ_6	0	32.95	0.02076	41.14	0.06676	47.07	0.1140	52.43

3.9 Conclusions

- The FORM and SORM formulations developed in the Chapter prove highly efficient and easy to implement in complicated reliability analysis. They can be used to solve practical engineering problems.

- Traditional numerical algorithms such as SQP method are often proposed for the general purpose. In reliability analysis they must be tailored to suit particular situation. This can save a lot of computational efforts.

- Among the optimisation methods in FORM, SQP method is the most robust one. It requires less iteration than other methods. However its efficiency decreases with increase of the variable size. For particular problem, the user must strike a balance between different methods.

- For the same limit state function, there may exist more than one minimum point. This can be tested simply by using different initial values in iteration.

- The nonlinearity of a problem is decided by the failure surface in *U* space. It is subject to the limit state function, the distribution of basic variables, and the correlation.

- The distribution of basic variable plays a very important part in reliability level. It must be appropriately chosen for particular problem.

- The reliability obtained by assuming basic variables are independent may be either conservative or unconservative. Correlation between basic variables should be taken into account whenever it is possible.

- In SORM, point fitting approach costs much less than curvature fitting. It is preferable in most cases, especially when the noise is unavoidable in calculating limit state function and its derivatives. The latter, however, is superior in some highly non-linear cases where the nth co-ordinate of fitting point is difficult to obtain without loss of probability information.

- Though Cai's and Koyluoglu's formulae are supposed to be useful when Breitung's and Tvedt's formulae are singular. They are more liable to fail especially when the

radii of main curvature are small. In practice, we can choose the one that has the highest precision.

References

1. Abdo, T., & Rackwitz R. (1990). A new beta-point algorithm for large time-invariant and time-variant reliability problems. *Proceedings of the 3rd WG 7.5 IFIP working conference*: *Reliability and optimization of structural systems*. Berkeley.

2. Ang, A. H. -S., & Tang, W. H. (1984). *Probability Concept in Engineering Planning and Design*. New York: Wiley.

3. Bleistin, N., & Handelsman, R. A. (1975). *Asymptotic expansions of integrals*. New York: Holt, Rinehart and Wilson.

4. Breitung, K. (1984) Asymptotic approximation for multinormal integrals. *Journal of Engineering Mechanics*, ASCE, 110(3), 357-366.

5. Breitung, K. (1989). Asymptotic approximation for probability integrals. *Probabilistic Engineering Mechanics*, 4(4), 187-190.

6. Cai, G., Q., & Elishakoff, I. (1994). Refined second-order reliability analysis. *Structural safety*, 14(4), 267-276.

7. Cornell, C. A. (1969). A probability-based structural code. *Journal of the American Concrete Institute*, 66(12), 974-985.

8. Der Kiureghian, A., ASCE, M., Lin, H., & Hwang, S. (1987). Second-order reliability approximations. *Journal of Engineering Mechanics*, ASCE, 113(8), 1208-1225.

9. Der Kiureghian, A., ASCE, M. & Liu, P. (1986). Structural reliability under incomplete probability information. *Journal of Engineering Mechanics*, ASCE, 112(1), 85-104.

10. Der Kiureghian, A., Zhang, Y., & Li, C.-C. (1994). Inverse reliability problem. *Journal of Engineering Mechanics*, ASCE, 120(5), 1154-1159.

11. Ditlevsen, O. (1981). Principle of normal tail approximation. *Journal of Engineering Mechanics Division*, ASCE, 107(6), 1091-1208.

12. Ditlevsen, O., & Madsen, H. O. (1996). *Structural reliability methods*, New York: John Wiley & Sons.

13. Dolinski, K. (1983). First order second moment approximation in reliability of systems: critical review and alternative approach. *Structural Safety*, 1(3), 211-231.

14. Fiessler, B., Rackwitz, R., & Neumann, H.-J. (1979). Quadratic limit states in structural reliability. *Journal of Engineering Mechanics Division*, ASCE, 105(4), 661-676.

15. Freudenthal, A. M. (1956). Safety and the probability of structural failure. *Transactions*, ASCE, 121, 1337-1397.

16. Hasofer, A. M., & Lind, N. C. (1974). An exact and invariant first order reliability format. *Journal of Engineering Mechanics*, ASCE, 100(EM1), 111-121.

17. Hasofer, A. M. (1974). Reliability index and failure probability. *Journal of Structural Mechanics*, 3(1), 25-27.

18. Helstrom, C. W. (1983). Comment: Distribution of quadratic forms in normal random variables-evaluation by numerical integration. *SIAM Journal on Scientific and Statistical Computing*, 4(2), 353-356.

19. Hohenbichler, M. & Rackwitz, R. (1986). Sensitivity and importance measures in structural reliability. *Civil Engineering Systems*, 3(4), 203-209.

20. Koyluoglu, H. U., & Nielsen, S. R. K. (1994). New approximations for SORM integrals. *Structural Safety*, 13(4), 235-246.

21. Liu, P. L., & Der Kiureghian, A. (1991). Optimization algorithms for structural reliability. *Structural safety*, 9(3), 161-177.

22. Luenberger, D. G. (1973). *Introduction to linear and nonlinear programming*, New York: Addison-Wesley.

23. Gill, P. E., Murray, W., & Wright, M. H. (1981) *Practical optimization*, London: Academic Press.

24. Bazaraa, M., Sherali, H. D., & Shetty, C. M. (1993). *Nonlinear programming: Theory and algorithm* (2nd ed.), New Hersey: John Wiley & Sons.

25. Madsen, H. O., Krenk, S. & Lind, N. C. (1986). *Methods of structural safety*, New Jersey: Prentice-Hall.

26. Madsen, H. O. (1988). Omission sensitivity factors. *Structural Safety*, 5(1), 35-45.

27. Melchers, R. E. (1999). *Structural reliability analysis and prediction* (2nd ed.). New York: John Wiley & Sons.

28. Maes., M. A. (1996). Ignorance factors using model expansion. *Journal of Engineering Mechanics*, ASCE, 122(1), 39-45.

29. Nataf, A. (1962). Determination des distribution de probabilites dont les marges sont donnes. *Comptes Rendus de l'Academie des Sciences*, 225, 42-43.

30. Racwitz, R. (1976) Practical probabilistic approach to design. *Bulletin* 112, Comite European du Beton.

31. Rice, S. O. (1980). Distribution of quadratic forms in normal random variables - evaluation by numerical integration. *SIAM Journal on Scientific and Statistical Computing*, 1(4), 438-448.

32. Rosenblatt, M., (1952). Remarks on a multivariate transformation. *Annals of Mathematics of Statistics*, 23(3), 470-472.

33. Sakamoto, J., Mori, Y., & Sekioka, T. (1997). Probabilistic analysis method using fast Fourier transform and its application. *Structural Safety*, 19(1), 21-36.

34. Schittkowski, K. (1980). Nonlinear programming codes: Information, tests, performance. *Lecture Notes in Economics and Mathematical Systems*, 183, Berlin: Spring-Verlag.

35. Schittkowski, K. (1981a). The nonlinear programming methods of Wilson, Han, and Powell with an augmented Lagrangian type line search function. Part 1: Convergence analysis. *Numerical Mathematics*, 38(1), 83-114.

36. Schittkowski, K. (1981b). The nonlinear programming methods of Wilson, Han, and Powell with an augmented Lagrangian type line search function. Part 2: An efficient implementation with linear least squares subproblems. *Numerical Mathematics*, 38(1), 115-127.

37. Schittkowski, K. (1983). On the convergence of a sequential quadratic programming method with an augmented Lagrangian line search function. *Mathematische Opertionsforschung und Statistik*, 14, 197-216.

38. Schittkowski, K. (1985). NLPQL - A Fortran subroutine solving constrained nonlinear programming problems. *Annals of Operations Research*, 5(1-4), 485-500.

39. Shinozuka, M. (1983). Basic analysis of structural safety. *Journal of Structural Engineering*, 109(3), 721-740.

40. Tvedt, L. (1984). Two second-order approximations to the failure probability. *Section on Structural Reliability*, Hovik: A/S Vertas Research.

41. Tvedt, L. (1985). On the probability content of a parabolic failure set in a space of independent standard normally distributed random variables. *Section on Structural Reliability*, Hovik: A/S Vertas Research.

42. Tvedt, L. (1990). The distribution of quadratic forms in normal space. *Journal of Engineering Mechanics*, ASCE, 116(6), 1183-1197.

43. Zhao, G. F. (1996). *Reliability theory and its applications for engineering structures*, Dalian: Dalian University of Technology Press.

Chapter 4 Monte Carlo Integration

4.1 Introduction

In the first and second order reliability analysis the failure surface is approximated by the first and second order expansion of the actual limit state function respectively. Tough this is usually sufficient in structural reliability analysis, the limit state could be poorly represented when the nonlinearity of the problem is high. Moreover, in FORM and SORM we need to know either the gradient or Hessian matrix of the limit state function. They are easy to obtain when the limit state function is simple and explicit. However, in many structural analyses only numerical methods are available. Sometimes even if the limit state function is known explicitly the expression of the corresponding derivatives may be too complicated as in crack propagation analysis. In these cases the derivatives can be calculated by differentiation method. But this often comes with severe loss of accuracy because of the well-known round-off problems. To attack all these problems we can resort to other integration techniques such as Monte Carlo simulation (MCS). It is preferred in highly non-linear cases where even SORM gives misleading results. For those problems that can be dealt with by FORM and SORM MCS can relieve us of the strenuous (sometimes intractable) sensitivity analysis while offering higher precision. The first part of this chapter is a brief review of Monte Carlo integration in reliability analysis. Then a new search based adaptive kernel method is proposed to improve at once the accuracy and the sampling efficiency. Finally a series of examples are designed to do justice to the proposed method.

4.2 Direct Monte Carlo method

In general the failure probability of a structure can be expressed as the following multiple integral

$$P_f = \int_D f(\boldsymbol{x}) d\boldsymbol{x} \tag{4.1}$$

where $\boldsymbol{x}$ is the random vector of basic variables, $f(\boldsymbol{x})$ is the joint p.d.f., and D is the failure region defined by $D = \{\boldsymbol{x} \mid G(\boldsymbol{x}) \le 0\}$. If the integration is performed in the whole space R^n, Eq. (1) can be rewritten as

$$P_f = E\{I[G(\boldsymbol{X})]\} = \int I[G(\boldsymbol{x})] f(\boldsymbol{x}) d\boldsymbol{x} \tag{4.2}$$

where $I[G(\boldsymbol{x})]$ is the indicator of failure. It is defined by

$$I[G(\boldsymbol{x})] = \begin{cases} 1 & G(\boldsymbol{x}) \le 0 \\ 0 & G(\boldsymbol{x}) > 0 \end{cases} \tag{4.3}$$

If N samples are generated from $f(\boldsymbol{x})$, the moment estimate of failure probability is

$$\hat{P}_f = \frac{1}{N} \sum_{i=1}^{N} I[G(\boldsymbol{x}_i)] = \frac{n_f}{N} \tag{4.4}$$

where n_f is the number of sample falling into the failure region. The mean value of $\hat{P}_f$ is

$$E(\hat{P}_f) = \sum_{i=1}^{n} \frac{1}{N} E[I(G(\boldsymbol{x}_i)] = P_f \tag{4.5}$$

and the variance is

$$Var(\hat{P}_f) = \sum_{i=1}^{N} \frac{1}{N^2} Var[I(G(\boldsymbol{x}_i)] = \frac{Var[I(G(\boldsymbol{x})]}{N} \tag{4.6}$$

Apprently, $\hat{P}_f$ here is an unbiased estimator of the actual failure probability. Shooman (1968) has suggested a confidence interval for n_f on the basis of central limit thereom and binomial assumption. Provided that $NP_f \ge 5$ and $P_f \le 0.5$, it can be written as

$$P\{-k[NP_f(1-P_f)]^{1/2} < n_f - NP_f < k[NP_f(1-P_f)]^{1/2}\} = C \tag{4.7}$$

where C is the confidence level, k is the corresponding upper fractile of standard normal distribution. Define error

$$\varepsilon = \frac{n_f - NP_f}{NP_f} \tag{4.8}$$

we have

$$|\varepsilon| < k\left(\frac{1-P_f}{NP_f}\right)^{1/2} \tag{4.9}$$

with confidence level C. Equation (7) can be used as a coarse estimate of the required sample size. Obviously, when the failure reliability is small a very large N is needed to obtain a rational estimate of P_f.

4.3 Importance sampling method

4.3.1 Theory of importance sampling

According to the above discussion, the only way to reduce the variance of $\hat{P}_f$ in direct Monte Carlo method is to increase the sample size N. This is an intolerable penalty when the failure probability is small and the failure function is costly to calculate. To relive the problem different variance reduction techniques have been developed. An overview of the various strategies was given by Rubinstein (1981). Among them importance sampling method is the most widely tested and adopted. This technique was introduced into structural reliability analysis first by Shinozuka (1983) and soon won great popularity.

The basic idea of importance sampling is to generate samples in such a way that they concentrate more in failure region D. The idea can be presented as

$$P_f = \int \frac{I[G(\boldsymbol{x})]f(\boldsymbol{x})}{h(\boldsymbol{x})} h(\boldsymbol{x})d\boldsymbol{x} \tag{4.10}$$

where $h(\boldsymbol{x})$ is the importance sampling density, from which the samples are drawn. An unbiased estimator of Eq. (10) is

$$\hat{P}_f = \frac{1}{N}\sum_{i=1}^{N}\frac{I[]f(\boldsymbol{x}_i)}{h(\boldsymbol{x}_i)} \tag{4.11}$$

with variance

$$Var(\hat{P}_f) = \frac{1}{N}\left\{\int\frac{I[G(\boldsymbol{x})]f^2(\boldsymbol{x})}{h^2(\boldsymbol{x})}h(\boldsymbol{x})d\boldsymbol{x} - P_f^2\right\} \tag{4.12}$$

The unbiased estimator of Eq. (12) is

$$\hat{Var}(\hat{P}_f) = \frac{1}{N-1}\left(\frac{1}{N}\sum_{i=1}^{n}\frac{I[G(\boldsymbol{x}_i)]f^2(\boldsymbol{x}_i)}{h^2(\boldsymbol{x}_i)} - \hat{P}_f^2\right) \tag{4.13}$$

It can be noticed that if we happen to choose

$$h(X) = \frac{I[G(\boldsymbol{x})]f(\boldsymbol{x})}{P_f} \tag{4.14}$$

the variance of $\hat{P}_f$ will be reduced to zero. This result seems useless since if we know P_f there is no point using Monte Carlo integration. However it hints that we can at least choose the sampling density in such a manner that its shape is similar to $f(\boldsymbol{x})$ in the failure region.

If we use $f(\boldsymbol{x})$ for $h(\boldsymbol{x})$ in Eq. (12), the variance of direct Monte Carlo simulation can be expressed as

$$Var(\hat{P}_f) = \frac{1}{N}(P_f - P_f^2) \tag{4.15}$$

The corresponding c.o.v is

$$\delta_{P_f} = \frac{\sqrt{Var(\hat{P}_f)}}{E(\hat{P}_f)} = \sqrt{\frac{1-P_f}{NP_f}} \tag{4.16}$$

or

$$N = \frac{1-P_f}{\delta_{P_f}^2 P_f} \tag{4.17}$$

This is another important estimate of required sample size in direct Monte Carlo. Assume $P_f = 1.0\times10^{-4}$, $\delta_{P_f} = 0.05$, N will be as large as 4,000,000. This greatly justifies the necessity of importance sampling technique.

4.3.2 Choice of random space

Following the pioneering work of Shinozuka (1983), dozens of importance sampling schemes have been proposed in structural reliability analysis. The differences between them reside in two aspects: The choice of space in which the simulation is carried out and the choice of appropriate sampling density.

Some researchers perform the Monte Carlo simulation in X space in order to avoid the transformation respecting non-normality and correlation. However more people prefer to carry it out in U space. In the author's point of view U space has many unbeatable advantages. First, the favorable mathematical properties of multivariate standard normal distribution make it possible to design and interpret an algorithm intuitively. Second it produces unbiased sample with respect to each variable and has greater sampling density in the region of most significance. Third different problem can be treated in a fairly uniform way regardless of the original distribution of basic variables and the correlation between them. Forth, when the correlated basic variables are neither normal nor lognormal the distribution transformation is unavoidable to calculate the joint pdf as in Nataf model.

Instead of Cartesian coordinate system polar coordinate system can be used in X or U space. Sprouting in the work of Deak (1980), this idea has been developed into a new class of Monte Carlo integration method termed directional simulation (Ditlevsen, 1986, 1988, 1989, 1990; Bjerager, 1988). This method is useful when there exists many a design point

or when the failure domain is specified by a union of failure modes. The drawback of directional simulation is that line search has to be performed to find the radius for each sample generated. Sometimes this makes it less robust or more time-consuming.

4.3.3 Choice of importance sampling density

The construction of sampling density is the central and very flexible part in Monte Carlo method. In favor of application, the distribution function should concentrate on the failure region, and most importantly the random numbers can be readily drawn from it. At first, Shinozuka (1983) used the uniform function defined over a hyper-rectangle. Another choice of the sampling density is the original joint pdf with mean vector shifted to the point of maximum likelihood (PML) (Harbitz, 1983; Ibrahim 1991). From an analytic point of view, Maes et al. (1993) constructed a series of sampling density satisfying Eq. (14). The method is called asymptotic sampling since it is based on the assumption that engineering systems are designed to a high standard of reliability. In implementation the gradients of the likelihood function and limit state function need to be calculated. Now more and more people tend to use multivariate normal distribution centered at PML (Schueller and Stix 1987). In this case, if the nonlinearity of the limit state function is not high, about 50% of the sampling point will be in the failure region. It is interesting to notice that this method can be applied to the integration of (3.143) in rotated space to improve the result of SORM (Hohenbichler and Rackwitz 1988). Typical among other sampling schemes is censored sampling suggested by Harbitz (1986) in U space and generalized by Melchers (1989). In this method sampling is prevented in the region that has no contribution to failure probability.

To further improve the sampling efficiency different adaptive Monte Carlo methods have been developed. The point is to make the most of the information available in the failure region. The knowledge of the failure region is usually obtained by initial sampling. Then it is used to update the sampling density in the succeeding importance resampling. The idea was proposed by Bucher (1988). In his method the requirement in Eq. (14) is satisfied in terms of the first and second moments. Based on the same methodology Karamchandani et al. (1989) introduce a method which yields good results dispensing with PML. However since these two methods begin with crude MCS, the initial sampling is very extensive when the failure probability is small. As an improvement, Melchers (1990) proposes a search-based adaptive formulation, in which the sampling density is kept shifting either to

the current minimum point in the safe region or to the current PML in the failure region. If performed in *U* space, this approach often gives an approximation of design point at the end of the initial sampling. In this method, a new estimator of failure probability is also used to take advantage of all the available sampling points. Melchers (1991) gives a comprehensive review of importance sampling method. A benchmark study about those popular methods can be found in the paper by Engelund and Rackwitz (1993).

In the methods above mentioned, the family of sampling density is specified. What changes in the adaptation is only the parameter. So they are actually a kind of parametric estimation method. To obtain more adaptability we have to allow certain flexibility in the distribution of observed data in the failure region. In other words the data should be given more rights to speak for themselves. The non-parametric density estimation methods can be used for this purpose. Among them kernel method is the most popular one with statisticians (Silverman, 1986; Wand and Jones, 1995; Bowman and Azzalini, 1997). This method is introduced into importance sampling in structural reliability analysis by Ang et al. (1989,1992). Its application in structural system reliability has been presented by Wang et al (1997). Lately, Au and Beck (1999) proposed a new adaptive version based on kernel method. They used Markov chain in the initial sampling as approximate samples from Eq. (14). This idea will be used as a prototype in the new approach proposed below.

4.3.4 Point of maximum likelihood

Usually PML can be identified by numerical maximization. In *U* space it happens to be the design point. Otherwise there is no exact correspondence between them and no unique maximum point may exist. Breitung (1991) has pointed out that PML and design point are asymptotically equivalent to each other. The search-based technique suggested by Melchers (1990) is useful to find an approximate PML, other techniques are needed though to prevent the samples from being trapped in the area of local maximum point.

4.4 Generation of posterior distribution by Markov chain Monte Carlo method

The optimum sampling density given by Eq. (14) is usually non-Gaussian, complex, and multivariate. Traditionally acceptance and rejection method (A-R) can be used to generate random number from such a distribution (Rubinstein 1981). Recently Markov chain Monte Carlo (MCMC) method has drawn more and more attention of the statisticians (Tierney,

1994; Gamerman, 1997). It proves powerful in posterior estimate and inference. The Metropolis-Hastings (M-H) algorithm is the most versatile method in MCMC. It was proposed first by Metropolis et al. (1953) and generalised by Hastings (1970). It will be used to fit the initial kernel estimate in section 4.6.

On a continuous state space the Markov chain can be defined as

$$P_r(\boldsymbol{\theta}^{(n+1)} \in \boldsymbol{A} \,|\, \boldsymbol{\theta}^{(n)} = \boldsymbol{x}, \boldsymbol{\theta}^{(n-1)} = \boldsymbol{x}_{n-1}, \ldots, \boldsymbol{\theta}^{(0)} = \boldsymbol{x}_0) = P_r(\boldsymbol{\theta}^{(n+1)} \in \boldsymbol{R} \,|\, \boldsymbol{\theta}^{(n)} = \boldsymbol{x}) \quad (4.18)$$

for $\boldsymbol{x} \in \boldsymbol{R}^d$, $\boldsymbol{A} \in \boldsymbol{\mathcal{B}}$, and $n \in T$, where $\boldsymbol{\mathcal{B}}$ is the Borel σ-field on $\boldsymbol{R}^d$, and T is a discrete parameter space. A Markov chain is called homogeneous when the conditional probabilities in (18) do not depend on the step n. In this case we can define a transition kernel as conditional distribution function, which represents probability of moving from $\boldsymbol{x}_n$ to a point in set A

$$P(\boldsymbol{x}_n, \boldsymbol{A}) = P_r(\boldsymbol{x}_{n+1} \in \boldsymbol{A} \,|\, \boldsymbol{x}_0, \ldots, \boldsymbol{x}_n) \quad (4.19)$$

The conditional transition probability over m steps is given by

$$P^m(\boldsymbol{x}_n, \boldsymbol{A}) = P_r(\boldsymbol{x}_{n+m} \in \boldsymbol{A} \,|\, \boldsymbol{x}_n) \quad (4.20)$$

It satisfies the Chapman-Kolmogorov equation

$$P^{n+m}(\boldsymbol{x}, \boldsymbol{A}) = \int_{R^d} P^n(\boldsymbol{x}, d\boldsymbol{y}) P^m(\boldsymbol{y}, \boldsymbol{A}), \qquad m, n \geq 0 \quad (4.21)$$

This is an integral in Lebesgue sense. Generally with initial distribution μ and transition function $P(\cdot,\cdot)$, a stationary Markov chain can be constructed as

$$\begin{aligned} &P_\mu(\boldsymbol{A}_0 \times \boldsymbol{A}_1 \times \cdots \times \boldsymbol{A}_n) = P_\mu(\boldsymbol{x}_0 \in \boldsymbol{A}_0, \cdots, \boldsymbol{x}_n \in \boldsymbol{A}_n) \\ &= \int_{\boldsymbol{A}_0} \mu(d\boldsymbol{x}_0) \int_{\boldsymbol{A}_1} P(\boldsymbol{x}_0, d\boldsymbol{x}_1) \cdots \int_{\boldsymbol{A}_{n-1}} P(\boldsymbol{x}_{n-2}, d\boldsymbol{x}_{n-1}) P(\boldsymbol{x}_{n-1}, \boldsymbol{A}_n) \end{aligned} \quad (4.22)$$

The invariant distribution π^* of a Markov chain is defined by

$$\pi^*(d\boldsymbol{y}) = \int_{R^d} P(\boldsymbol{x}, d\boldsymbol{y})\pi^*(d\boldsymbol{x}) \tag{4.23}$$

One of the main tasks in Markov chain theory is to study under which condition the following limit can be achieved:

$$\lim_{n\to\infty} P^n(\boldsymbol{x}, \boldsymbol{A}) = \pi^*(\boldsymbol{A}) \tag{4.24}$$

This problem is turned around in Marchov chain Monte Carlo (MCMC) method. Now what we are concerned about is how to construct a transition kernel with target distribution π^*. An answer to this question is as follows. Assume transition kernel can be expressed as

$$P(\boldsymbol{x}, d\boldsymbol{y}) = p(\boldsymbol{x}, \boldsymbol{y})\mu(d\boldsymbol{y}) + r(\boldsymbol{x})\delta_x(d\boldsymbol{y}) \tag{4.25}$$

where density $p(\boldsymbol{x},\boldsymbol{x}) = 0$, $\delta_x(d\boldsymbol{y}) = 1$ if $\boldsymbol{x} \in d\boldsymbol{y}$ and 0 otherwise and

$$r(\boldsymbol{x}) = 1 - \int_{R^d} p(\boldsymbol{x}, \boldsymbol{y})d\boldsymbol{y} \tag{4.26}$$

Obvious r is the probability that the chain remains at $\boldsymbol{x}$. It has been proved that if p satisfies the reversibility condition

$$\pi(\boldsymbol{x})p(\boldsymbol{x}, \boldsymbol{y}) = \pi(\boldsymbol{y})p(\boldsymbol{y}, \boldsymbol{x}) \tag{4.27}$$

then π is the invariant density of P (Tierney 1994). In Metropolis-Hastings (M-H) method the density p is constructed as

$$p(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} q(\boldsymbol{x}, \boldsymbol{y})\alpha(\boldsymbol{x}, \boldsymbol{y}), & \text{if } \boldsymbol{x} \neq \boldsymbol{y} \\ 0, & \text{if } \boldsymbol{x} = \boldsymbol{y} \end{cases} \tag{4.28}$$

where q is the candidate density, and α is the probability of move defined by

$$\alpha(\boldsymbol{x},\boldsymbol{y}) = \begin{cases} \min\left\{\dfrac{\pi(\boldsymbol{y})q(\boldsymbol{y},\boldsymbol{x})}{\pi(\boldsymbol{x})q(\boldsymbol{x},\boldsymbol{y})},1\right\}, & \text{if } \pi(\boldsymbol{x})q(\boldsymbol{x},\boldsymbol{y}) > 0 \\ 1, & \text{if } \pi(\boldsymbol{x})q(\boldsymbol{x},\boldsymbol{y}) = 0 \end{cases} \tag{4.29}$$

The candidate density is usually selected from a family of distribution with location and scales to be specified. However a complete theory respecting how to choose an appropriate candidate is still wanted. Among the existing practical choices, the multivariable normal and the multivariable-*t* density are the most preferable. Since they are symmetric, equation (29) will be reduced to

$$\alpha(\boldsymbol{x},\boldsymbol{y}) = \min\left\{\frac{\pi(\boldsymbol{y})}{\pi(\boldsymbol{x})},1\right\} \tag{4.30}$$

The scale of the candidate density affects the behaviour of the chain in two aspects. In one way it controls the region that is covered by the chain and in the other the acceptance rate, namely the rate at which a move to a new point is made. Hence a large scale means more important region will be covered. From Eq. (30) we know the new step could be either uphill or downhill. To some extent, this can prevent the chain from being trapped when there is more than one design point in this region. However a relatively large scale will definitely increase the number of repeated sample and decrease the acceptance rate. So a balance has to be struck here. Chib and Breenberg (1995) have given a brief review in determining the candidate density along with an intuitive interpretation to Eq. (28) and Eq. (29).

If we perform the Monte Carlo integration in *U* space the multivariate normal distribution is obviously a nice choice for the candidate density. In such a context the posterior density in Eq. (14) can be generated as follows:

1) Initialize the starting point $\boldsymbol{x}^{(0)}$ in the failure region, set $j = 0$.

2) Generate $\boldsymbol{y}$ from q and u from a uniform generator over (0,1).

3) If $u \le \frac{f(\boldsymbol{y})}{f(\boldsymbol{x}^{(j)})}$, set $\boldsymbol{x}^{(j+1)} = \boldsymbol{y}$, otherwise set $\boldsymbol{x}^{(j+1)} = \boldsymbol{x}^{(j)}$.

4) 4. Set $j = j + 1$, repeat step 2-3 until the whole Markov chain is generated.

4.5 Kernel method in estimation of the importance sampling density

4.5.1 Kernel density estimator

In one dimensional random space, once a sample set $D_m = \{x_1, x_2, \ldots, x_m\}$ is obtained, the p.d.f of random variable x can be estimated by

$$\hat{f}(x; D_n, b) = \frac{1}{m}\sum_{i=1}^{m} I(x; y_i, b) \tag{4.31}$$

where $I(x; y, b)$ is the indicator function on the window $[y-b, y+b]$. This is the well-known histogram approach. It suffers from the following criticisms:

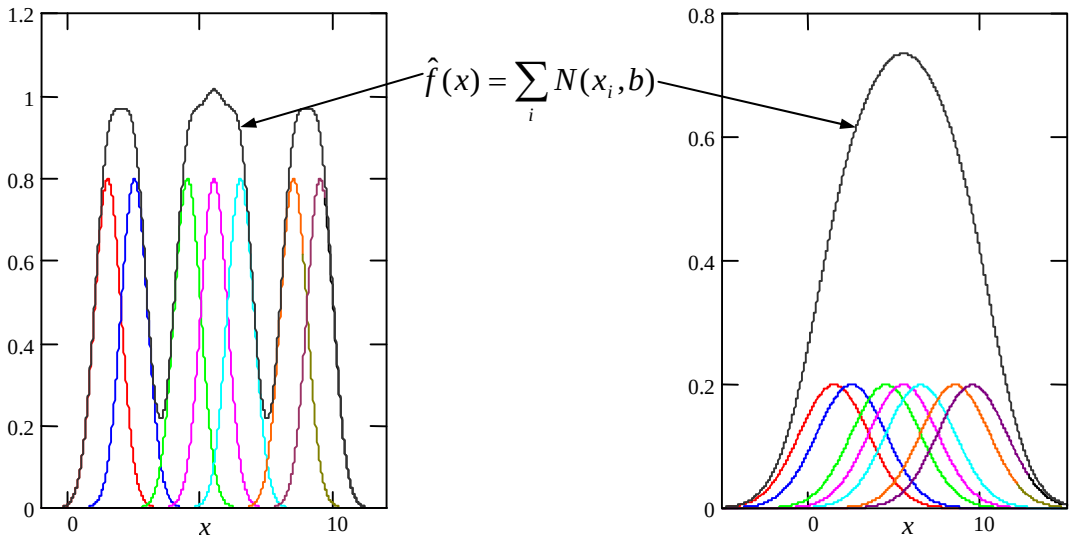
- Information loss when replacing $\{x_1, x_2, \ldots, x_m\}$by $\{y_1, y_2, \ldots, y_m\}$.
- The underlying function is assumed to be smooth but the estimator is apparently not.
- The behavior of the estimator is dependent on the choice of windows and, in particular, the window width.

The first two problems were removed by the work of Roseblatt (1956), Whittle (1958) and Paren (1962) by introducing estimator

$$\hat{f}(x; D_m, b) = \frac{1}{m}\sum_{i=1}^{m} K(x; x_i, b) \tag{4.32}$$

Here K itself is a probability density termed kernel function, b is named window width or smoothing parameter. It is often convenient to choose normal density for K. Other popular choices for K are the T-distributions or more refined forms like split T-distributions

(Geweke 1989). The asymptotic behavior of the kernel density is relatively insensitive to the form of kernel function (Silverman 1986). However *b* must be chosen with care, since it affects the manner in which the probability associated with each observation is spread (see Fig.4.1). A large value of *b* tends to overspread the kernel density, while a small value may cause spurious noise at the tail of the distribution where the samples are sparsely populated. Often a minimization with respect to some objective function is needed to determine *b*. It can be noticed that the smoothness of *K* is now inherited by $\hat{f}$ with a kernel function centered directly over each observation. Actually Eq. (32) can be looked on as a generalization of Eq. (31). Other types of estimator exist such as those based on orthogonal series (Fryer, 1977). But no evidence shows that one estimator is superior to others and the choice of a smoothing factor analogous to the window width *b* is always required. A general introduction to the subject of density estimation is the review paper by Fryer (1977) and by Wertz and Schneider (1979).



(a) *b* = 0.5 (b) *b* = 2.0

Figure 4.1 Effect of smoothing factor on the kernel mixture

In general, the kernel density for a random vector can be constructed as

$$\hat{f}(\boldsymbol{x}) = \frac{1}{m}\sum_{i=1}^{m} K(\boldsymbol{x}; \boldsymbol{x}_i, b, \lambda_i, \boldsymbol{C}_i) \tag{4.33}$$

where C_i is the local covariance matrix, b is the global smoothing factor, and λ_i is the local smoothing factor. In Eq. (33) the multivariate normal distribution can be used as the kernel function. The local covariance matrix C_i has to be evaluated particularly for each kernel function. Along with the local smoothing factor λ_i it allows us to control the envelope of the kernel density in a more delicate way. This is important when the target distribution is highly nonconvex or there are strong nonlinear relationships between random variables (Givens and Raftery, 1996).

4.5.2 Construction of kernel density in Monte Carlo simulation

The optimal sampling density h is posterior. In failure region it is only known to a constant of proportionality, namely $f(\boldsymbol{x}) \propto h(x)$. Consider the problem of evaluating

$$E[t(\boldsymbol{x})] = \int t(\boldsymbol{x})h(\boldsymbol{x})d\boldsymbol{x} \tag{4.34}$$

It is equivalent to

$$I = \frac{\int I[G(\boldsymbol{x})]t(\boldsymbol{x})\, f(\boldsymbol{x})d\boldsymbol{x}}{\int I[G(\boldsymbol{x})] f(\boldsymbol{x})} \tag{4.35}$$

which can be evaluated by

$$\hat{I} = \frac{\sum_{i=1}^{n} t(\boldsymbol{x}_i)\, f(\boldsymbol{x}_i)/p(\boldsymbol{x}_i)}{\sum_{i=1}^{n} f(\boldsymbol{x}_i)/p(\boldsymbol{x}_i)} = \sum_{i=1}^{n} w_i t(\boldsymbol{x}_i) \tag{4.36}$$

where

$$w_i = \frac{f(\boldsymbol{x}_i)/p(\boldsymbol{x}_i)}{\sum_{i=1}^{n} f(\boldsymbol{x}_i)/p(\boldsymbol{x}_i)} \tag{4.37}$$

Here p is an importance sampling density. It dominates the tail of f. As a further generalization of Eq. (33) the kernel density can be constructed by the following integral

$$E[K(\boldsymbol{x};\boldsymbol{\theta},b,\lambda,\boldsymbol{C})] = \int K(\boldsymbol{x};\boldsymbol{\theta},b,\lambda(\boldsymbol{\theta}),\boldsymbol{C}(\boldsymbol{\theta}))h(\boldsymbol{\theta})d\boldsymbol{\theta} \tag{4.38}$$

The corresponding estimator is

$$\hat{h}(\boldsymbol{x}) = \sum_{i=1}^{m} w_i K(\boldsymbol{x};\boldsymbol{x}_i,b,\lambda_i,\boldsymbol{C}_i) \tag{4.39}$$

where $\boldsymbol{x}_i$ are from a sample set of Eq. (14), say obtained by MCMC method. If we use a global covariance structure for the kernel density, it can be estimated by

$$\hat{\boldsymbol{C}} = \hat{\boldsymbol{C}}_i = \sum_{i=1}^{m} w_i(\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})^T \tag{4.40}$$

It is a Monte Carlo estimate. Otherwise $\boldsymbol{C}_i$ have to be estimated individually (Givens, 1995). This will increase the computational cost dramatically if the number of kernel component is large. Suppose the design point is unique, a global covariance estimate is usually enough. In this case, the local behavior of the kernel density can be adjusted only by the local smoothing factor λ_i. Otherwise we can assign the same covariance structure for each importance region.

4.5.3 Choosing the optimum global smoothing factor

4.5.3.1 MISE and maximum likelihood criteria

To measure the error of the constructed kernel density appropriate criteria need to be specified. The optimum global smoothing factor is the one that makes the error minimum. Apparently different criterion will lead to different answer. The most widely used criterion is the integrated squared error (ISE) given by

$$ISE(\hat{h}) = \int (\hat{h} - h)^2 d\boldsymbol{x} = \int (\hat{h}^2 - 2\hat{h}h + h^2) d\boldsymbol{x} \tag{4.41}$$

It is a measure over the whole real space. Sometimes the mean integrated squared error (MISE) is used, which is

$$MISE(\hat{h}) = E[ISE(\hat{h})] = E\int (\hat{h} - h)^2 \, d\boldsymbol{x} \tag{4.42}$$

Since MISE takes into account other possible data set from the target density *h*, it is more appropriate.

The error criteria can be defined according to specific purpose. However it is often possible to find their counterparts in statistics. One such criterion is given by Ang G. L. et al. (1989,1992) in structural reliability analysis. It is designed to minimise the cov of the estimate of failure probability. Later we will prove it is actually equivalent to the maximum likelihood criterion.

One problem with Eq. (41) and Eq. (42) is that they are too complicated for analytical analysis. Usually their second order Taylor approximations are used. The corresponding criteria thus deduced are called asymptotic MSE and MISE, denoted by AMSE and AMISE respectively. Suppose the target is *d*-variate $N(\boldsymbol{\mu},\boldsymbol{H})$ density the kernel function is unit *d*-variate normal density with equal weight factor, the global bandwidth matrix that minimises the AMISE will be (Wand and Jones, 1995)

$$\boldsymbol{C}_{\text{AMISE}} = \left(\frac{4}{n(d+2)} \right)^{\frac{2}{d+4}} \boldsymbol{H} \tag{4.43}$$

In fact only few cases can be dealt with analytically. Numerical methods are often used to minimise Eq. (41) or Eq. (42) with respect to *b*. This involves evaluation of Eq. (41) and Eq. (42). If a Gaussian kernel function is adopted, they can be estimated effectively by cross-validation technique. Two new estimators will be proposed below in terms of the general format given by Eq. (39).

4.5.3.2 Least square cross-validation (LSCV)

Since the last term of Eq. (41) has nothing to do with $\hat{h}$, the optimum window width can be chosen by minimising

$$ISE(\hat{h}) - \int h^2 d\boldsymbol{x} = \int \hat{h}^2 d\boldsymbol{x} - 2\int \hat{h} h d\boldsymbol{x} \tag{4.44}$$

Construct

$$\hat{h}_{-i}(\boldsymbol{x}) = \sum_{i \neq j} w_j K(\boldsymbol{x}; \boldsymbol{x}_j, b, \lambda_j, \boldsymbol{C}_j) \tag{4.45}$$

where

$$K(\boldsymbol{x}; \boldsymbol{x}_j, b, \lambda_{j,} \boldsymbol{C}_j) = \frac{1}{(b\lambda_j)^d \sqrt{(2\pi)^d |\boldsymbol{C}_j|}} \exp\left[\frac{-(\boldsymbol{x} - \boldsymbol{x}_j)^T \boldsymbol{C}_j^{-1} (\boldsymbol{x} - \boldsymbol{x}_j)}{2b^2 \lambda_j^2} \right] \tag{4.46}$$

An unbiased estimator of Eq. (44) can be defined as

$$LSCV(b) = \int \hat{h}^2 d\boldsymbol{x} - 2(m-1)^{-1} \sum_i \hat{h}_{-i}(\boldsymbol{x}_i) \tag{4.47}$$

The first term of Eq. (47) can be estimated by the convolution of normal kernel

$$\begin{aligned} &\int \hat{h}(\boldsymbol{x})^2 d\boldsymbol{x} = \int \sum_i w_i K(\boldsymbol{x}; \boldsymbol{x}_i, b, \lambda_i, \boldsymbol{C}_i) \sum_j w_j K(\boldsymbol{x}; \boldsymbol{x}_j, b, \lambda_j, \boldsymbol{C}_j) d\boldsymbol{x} \\ &= \sum_i \sum_j w_i w_j \int K(\boldsymbol{x}; \boldsymbol{x}_i, b, \lambda_i, \boldsymbol{C}_i) K(\boldsymbol{x}; \boldsymbol{x}_j, b, \lambda_j, \boldsymbol{C}_j) d\boldsymbol{x} \\ &= \sum_i \sum_j w_i w_j K^{(2)}(\boldsymbol{x}_i, \boldsymbol{x}_j; b, \lambda_i, \lambda_{,j} \boldsymbol{C}_i, \boldsymbol{C}_j) \end{aligned} \tag{4.48}$$

where

$$\begin{aligned} &K^{(2)}(\boldsymbol{x}_i, \boldsymbol{x}_j; b, \lambda_i, \lambda_j, \boldsymbol{C}_i, \boldsymbol{C}_j) \\ &= \frac{1}{b^d \sqrt{(2\pi)^d |(\lambda_i^2 \boldsymbol{C}_i + \lambda_j^2 \boldsymbol{C}_j)|}} \exp\left[\frac{-(\boldsymbol{x}_i - \boldsymbol{x}_j)^T (\lambda_i^2 \boldsymbol{C}_i + \lambda_j^2 \boldsymbol{C}_j)^{-1} (\boldsymbol{x}_i - \boldsymbol{x}_j)}{2b^2} \right] \end{aligned} \tag{4.49}$$

A special case of Eq. (45) is

$$\hat{h}_{-i}(\boldsymbol{x}) = (m-1)^{-1}\sum_{i\neq j} K(\boldsymbol{x};\boldsymbol{x}_j, b, \boldsymbol{C}) \tag{4.50}$$

Consequently Eq. (47) becomes

$$LSCV(b) = \int \hat{h}^2 d\boldsymbol{x} - 2m^{-1}\sum_i \hat{h}_{-i}(\boldsymbol{x}_i) \tag{4.51}$$

The density estimator given by Eq. (50) is often called "leave-one-out" density estimator. The way we construct h_{-i} it is referred to as cross-validation technique, which uses one part of the sample to obtain information about another part. Sometimes LSCV may have more than one minimum point. It is suggested that the largest local minimiser should be used instead of the global one (Wand and Jones, 1995). Intuitively speaking the kernel density thus formed will have a longer tail and cover more importance region. However, many studies have shown that the theoretical and practical performance of LSCV is somewhat disappointing (Park and Marron, 1990). It is sensitive to the variation of sample (Hall and Marron, 1987). So care must be taken here and experience always has a big say. If possible we had better use other bandwidth selectors as a cross-check.

4.5.3.3 Maximum Likelihood Cross-validation (MLCV)

The central task in importance sampling is to reduce the variance in estimating the failure probability. Naturally we can choose a smoothing factor to minimise Eq. (12). Substituting Eq. (14) into Eq. (12) we have

$$Var(\hat{P}_f) = \frac{1}{N}\left\{ P_f \int \frac{I[G(\boldsymbol{x})]f(\boldsymbol{x})}{\hat{h}(\boldsymbol{x})} h(\boldsymbol{x}) d\boldsymbol{x} - P_f^2 \right\} \tag{4.52}$$

The minimization of Eq. (50) is equivalent to

$$Min\left[Var(\hat{P}_f)\right] = Min\left\{ \int \frac{I[G(\boldsymbol{x})]f(\boldsymbol{x})}{\hat{h}(\boldsymbol{x})} h(\boldsymbol{x}) d\boldsymbol{x} \right\} \tag{4.53}$$

which is in turn equivalent to

$$Min\left[Var(\hat{P}_f)\right] = Min\left\{\int I[G(\boldsymbol{x})]\log\left[\frac{f(\boldsymbol{x})}{\hat{h}(\boldsymbol{x})}\right]h(\boldsymbol{x})d\boldsymbol{x}\right\} \tag{4.54}$$

Equation (54) is in nature the Kullback-Leibler information distance. Since

$$\begin{aligned}&\int I[G(\boldsymbol{x})]\log\left[\frac{f(\boldsymbol{x})}{\hat{h}(\boldsymbol{x})}\right]h(\boldsymbol{x})d\boldsymbol{x}\\ &= \int I[G(\boldsymbol{x})]\log[f(\boldsymbol{x})]h(\boldsymbol{x})d\boldsymbol{x} - \int I[G(\boldsymbol{x})]\log[\hat{h}(\boldsymbol{x})]h(\boldsymbol{x})d\boldsymbol{x}\end{aligned} \tag{4.55}$$

and the first term in Eq. (55) has nothing to do with the window width, we obtain

$$Min[Var(\hat{P}_f)] = Max\{\int I[G(\boldsymbol{x})]\log[\hat{h}(\boldsymbol{x})]h(\boldsymbol{x})d\boldsymbol{x}\} \tag{4.56}$$

This is in fact the well-known problem of maximum likelihood. The cost function can be estimated by the cross-validation

$$\int I[G(\boldsymbol{x})]\log[\hat{h}(\boldsymbol{x})]h(\boldsymbol{x})d\boldsymbol{x} \approx \sum_i w_i \log[\hat{h}_{-i}(\boldsymbol{x}_i)] \tag{4.57}$$

If Eq. (53) is used the estimator will be

$$MLCV(b) = \int I[G(\boldsymbol{x})]\log\left[\frac{f(\boldsymbol{x})}{\hat{h}(\boldsymbol{x})}\right]h(\boldsymbol{x})d\boldsymbol{x} \approx \sum_i w_i \log\left[\frac{f(\boldsymbol{x}_i)}{\hat{h}_{-i}(\boldsymbol{x}_i)}\right] \tag{4.58}$$

The estimators in Eq. (57) and Eq. (58) are asymptotically unbiased. A similar estimator has been suggested by Ang et al. (1989). But here it is put in a more general format with a brand new interpretation.

4.5.4 Determination of the local smoothing factor

Traditionally local bandwidth is defined as the distance from $\boldsymbol{x}_i$ to its kth nearest point in the data set. The parameter k decides how responsive the window width choice will be to

the local detail. The calculation is rather tedious though some fast algorithms are available such as those proposed by Friedman et al. (1975,1977). Now people turn to more intuitive methods. Assuming a pilot kernel estimate, $\hat{h}$, is obtained, the local smoothing factor can be determined by

$$\lambda_i = \left\{\hat{h}(\boldsymbol{x}_i)/g\right\}^{-\alpha} \tag{4.59}$$

where $0 \le \alpha \le 1$ is the sensitivity factor, and g is the geometric mean of $\hat{h}(\boldsymbol{x}_i)$:

$$\log g = m^{-1}\sum_i \log \hat{h}(\boldsymbol{x}_i) \tag{4.60}$$

Breiman et al. (1977) and Abramson (1982) have shown that λ_i is not sensitive to the pilot estimate. So a straightforward estimate would be a fixed kernel estimate with bandwidth based on multivariate normal distribution as in Eq. (43). Besides, particular smoothness is not needed in pilot estimate. All this allows us some flexibility in choosing the pilot kernel function. Apparently those that can be fast evaluated are preferred. Strictly speaking, the factor g^{α} is not necessary, since its effect has been represented by b. However as a reward it makes the smoothing factor dimensionless and the geometric mean of λ_i equal to one. Silverman (1986) pointed out that good results can be expected if the bandwidth b is used for both pilot estimate and the final estimate. This will involve an iteration procedure.

As to the sensitivity factor α, the larger it is the more sensitive the method will be to the variation of the pilot density, and the more difference there will be between bandwidth used in different parts of the sample. When $\alpha = 0$, the method reduces to fixed-width kernel estimator. On the other hand the choice of $\alpha = 1/d$ will ensure that the number of observations 'caught' by the scaled kernel will be approximately the same in all parts of the density (Breiman et al., 1977). Practical experience reported by Abramson (1982) among others suggests that the choice of $\alpha = 1/2$ gives good results. This gives rise to an estimate whose bias is of smaller order than that of the fixed-width estimate.

4.6 Adaptive importance sampling scheme

4.6.1 Adaptive kernel formulation

The accuracy of Monte Carlo simulation depends not only on how much information is accessible in the importance region but more importantly how the information is obtained and used. An ill used sample set of large size can not ensure a better result than that from a smaller sample set of superior structure. In adaptive importance sampling the knowledge of the importance region is accumulated in a progressive way. First a set of pilot samplings is carried out to fit the sampling density. A coarse estimate of failure probability can also be calculated in this stage. The importance re-sampling is than performed to evaluate the final reliability level. Based on the discussion hitherto, an adaptive approach is proposed below in U space with some new characteristics:

Algorithm 4.1

1) Search for an approximate PML $\boldsymbol{x}_0$ in the failure region;

2) Generate a Markov Chain $\boldsymbol{\theta}_0$ of length M_0, with target density in Eq. (14) and starting point $\boldsymbol{x}_0$,

3) Initialize weights $\boldsymbol{w}_0$, calculate covariance structure $\boldsymbol{C}_0$ by Eq. (40), and evaluate global smoothing factor b_0 by Eq. (43);

4) Fit the kernel estimator by

$$\hat{h}_0(\boldsymbol{x}) = \sum_{i=1}^{M_0} w_i K(\boldsymbol{x}; \boldsymbol{\theta}_0, b_0, \boldsymbol{C}_0) \tag{4.61}$$

as the first approximation to $\hat{h}$;

5) Set $j = 1$;

6) Draw a sample set of size M_j, denoted by $\boldsymbol{\theta}_j$ from $\hat{h}_{j-1}$, and calculate the weights $\boldsymbol{w}_j$ by Eq. (37). Update $\boldsymbol{C}_j$ and b_j accordingly;

7) Obtain a new kernel estimator $\hat{h}_j$ by Eq. (61) with updated parameters;

8) Evaluate the entropy of importance sampling weights relative to uniformity by

$$-\sum_{i=1}^{M_j} w_i \frac{\log w_i}{\log(M_j)} \tag{4.62}$$

9) Set $j = j+1$ and repeat step 6-8 k times until the entropy is close to one, usually we can take $k = 3$;

10) Collapse the kernel component from M_k to an appropriate number M_r;

11) Calculate local smoothing factor by Eq. (59);

12) Optimize the global smoothing factor by LSCV or MLCV;

13) Perform the importance sampling integration in Eq. (10) using the optimized kernel estimate $\hat{h}_{opt}$.

For the sake of simplicity, algorithm 4.1 is designed on the assumption that there is only one design point. The same covariance structure is used for all the kernel components. Besides in pilot sampling stage, the global bandwidth is approximated by the analytical result from the standard normal distribution. We believe that in U space it is a reasonable guess. Since all the discussion in section 4.5 are in terms of a general format. Algorithm 4.1 can be readily extended to more complicated cases. The technique described in the paper by West (1993) is a precursor of the proposed method.

4.6.2 Further comments on practical computation

4.6.2.1 The approximation of PML

The PML can be found by FORM or by some maximisation methods. But this will make the proposed method less independent. In fact the starting point for Markov Chain needs not to be exactly the PML itself. It can be any point in the importance region. The uphill method proposed by Melchers (1990) is an effective way to find such a point. The method is summarised as follows:

Algorithm 4.2

1) Choose a starting point in random space $\boldsymbol{x}^{(k)}$, (k = 1), usually the mean point of the basic variables;

2) Choose a sampling density ${}_k h(\boldsymbol{x})$ with mean vector $\boldsymbol{x}^{(k)}$ and covariance structure $b\boldsymbol{C}$;

3) Draw a sample from ${}_k h(\boldsymbol{x})$, denote it by ${}_k \boldsymbol{v}_j$;

4) If ${}_k \boldsymbol{v}_j \in D$, the failure region, calculate the joint pdf of the basic variables $f(\boldsymbol{x})$ and find the present maximum value, denote the corresponding point as $\boldsymbol{x}^*$;

5) Otherwise calculate the value of limit state function $G(\boldsymbol{x})$, find the present minimum value, denote the corresponding point as $\boldsymbol{x}^{**}$;

6) If ${}_k \boldsymbol{v}_j \in D$, update the sampling density ${}_k h(\boldsymbol{x})$ by shifting its mean vector to $\boldsymbol{x}^*$. Otherwise $\boldsymbol{x}^{**}$ is used as the mean vector;

7) Once a failure point is found, the bandwidth b can be reduced appropriately;

8) Set $k = k + 1$, repeat 3-7 M times;

9) The final $\boldsymbol{x}^*$ can be regarded as an approximation of PML.

In U space, the multivariate standard normal function is a good choice for sampling function in algorithm 4.2. The variances of the basic variables can be used as the diagonal elements of $\boldsymbol{C}$. In fact we can use the sampling density ${}_M h(\boldsymbol{x})$ to evaluate the failure probability if a resampling is carried out.

4.6.2.2 Entropy measure of refinement

The entropy in Eq. (62) is a measure of variability in the importance sampling weights. It provides rough and very informal guidelines in kernel density refinement. It is helpful to guide successive choices of sample sizes and successive kernel smoothing factors. As the successive importance sampling functions approach the target function h, the weight distributions will tend to be uniformity, and the entropy measure approaches 1. The reasoning behind it is that

$$\log M - \sum_i w_i \log w_i \tag{4.63}$$

is actually the Monte Carlo estimate of the Kullback-Leibler divergence

$$\int \log\left(\frac{\hat{h}(\boldsymbol{x})}{h(\boldsymbol{x})}\right) h(\boldsymbol{x}) d\boldsymbol{x} \tag{4.64}$$

4.6.2.3 Collapsing mixtures

In multi-dimensional problems, thousands of kernel components are often needed to fit a reliable kernel estimate. After the adaptive iteration some kernel components may be positioned so densely that the effect of any of them can be replaced by that of their nearest neighbours. In other words the absence of those redundant components is insignificantly different to all practical purposes like the calculation of the probability. Out of this consideration we might as well reduce the kernel estimates to mixtures of much smaller numbers of components to save computational costs in the following simulation. One way leading to the reduction of this redundancy is the 'collapsing', or 'clustering' techniques. As a typical approach the method proposed by West (1993) is given below:

Algorithm 4.3

1) Set $r = n$, the current number of component mixture, choose $k < n$, the number of components for the reduced mixture;

2) Sort the r values of x_j in θ in order of increasing values of weights w_j, thus x_1 corresponds to the component with the smallest weight;

3) Find the index i($i = 1,\ldots,r$) such that x_i is the nearest neighbor of x_1, and reduce the set θ to set of size r –1by removing component 1 and i and inserting average values

$$\bar{\boldsymbol{x}} = \frac{w_1 \boldsymbol{x}_1 + w_i \boldsymbol{x}_i}{w_1 + w_i} \tag{4.65}$$

4) Set $r = r - 1$, and repeat step 2-4 until $r = k$.

More general issues in the approximation of mixtures can be referred to the book by West and Harrison (1997).

4.7 Examples

4.7.1 Example 1

Assume the limit state function is

$$G = \mp\sum_{i=1}^{n} X_i \pm C \tag{4.66}$$

where X_i, $i = 1, 2,\ldots,n$ are independent and exponentially distributed with the parameter λ (Engelund and Rackwitz,1993). The limit state function is highly non-linear in the U space:

$$G = \pm\sum_{i=1}^{n} \frac{\ln[\Phi(-U_i)]}{\lambda} \pm C \tag{4.67}$$

The exact failure probability can be calculated by gamma distribution $F_G(C; n, \lambda)$. This example is used to test the capacity and precision of the proposed method. It is calculated for $P_f = 1.0\times10^{-6}$, $n = 10$, and $\lambda = 1.0$.

Let us consider the first limit function with $C = 32.710$. It has positive main curvatures in the standard normal space $\kappa_i = 0.127$, corresponding to a concave failure set. First the effect of the sample size in pre-sampling stage is studied within three adaptive iterations. The sample size N_p needed in pre-sampling can be expressed as

$$N_p = N_1 + N_2 + N_3 \tag{4.68}$$

where N_1 is the sample size needed to search for an approximate PML, N_2 is the sample size used to generate Markov chain, N_3 is the sample size required in adaptive iteration. N_3 can be further put as

$$N_3 = kN_s + \frac{k(k-1)}{2} N_c \tag{4.69}$$

where k is the number of adaptive iteration, N_s is the starting sapling size, and N_c is the sample increment after each iteration. Initially we set N_1=1000, N_2 = N_s = 3000, and N_c =1500. The effect of N_2 and N_s is studied in turn. Each time the values of other parameters remain unchanged. The results are given in Fig.4.2 and Fig.4.3. As is shown, the uniformity entropy of the density estimate increases with the increase of N_2 and N_3. However, N_3 has more contribution to the precision, especially when the Markov chain has converged after a certain sampling.

After three iterations according to the default values given above we have obtained 5282 kernel components. Before we start the resampling procedure, the number of component is reduced to 2000 by collapsing techniques in 4.6.2.3. The global window width is estimated by Eq. (43), LSCV, and MLCV respectively. The simulated results are shown in Fig.4.4 with both global and local scaling, and in Fig. 4.5 with only global scaling. It can be seen that the introduction of local scaling factor in the present problem does not improve the sampling function, instead it makes the results more sensitive to the optimisation scheme. This is because the tail of the fitted sampling function is over smoothed. In both cases the analytical global width from Eq. (43) leads to very good estimate of the probability of failure. This justifies the implementation of the algorithm in standard normal space and increases the computational efficiency. Fig.4.6 presents the effect of the component number in the collapsed kernel estimate. Here only global window width is used and it is optimised by MLCV. From Fig.4.6 we can see that the behaviour of the sampling function by kernel estimate is dominated by those kernel components with higher weighting coefficient. An appropriately collapsed sampling function will reduce the computational time without much loss of accuracy. However the extent to which the kernel estimate should be collapsed depends on specific problems.

Now let us turn to the second failure function with C = 1.277. It has negative main curvatures in the standard normal space κ_i = 0.444, corresponding to a convex failure set. The results of pre-sampling are given in Fig.4.7 respectively, with N_1 = 1000, N_2 = 1000, N_s = 1500, and N_c = 1500. After pre-sampling we obtain 1660 kernel components in the

estimated sampling function. In the following importance resampling a sampling function with 1000 kernel components is used by collapsing technique, the results are shown in Fig.4.8. Clearly, while the sample size we use in this case is much less than the previous one, the accuracy achieved is a lot higher. This reveals that the proposed method is sensitive to the convexity of the problem.

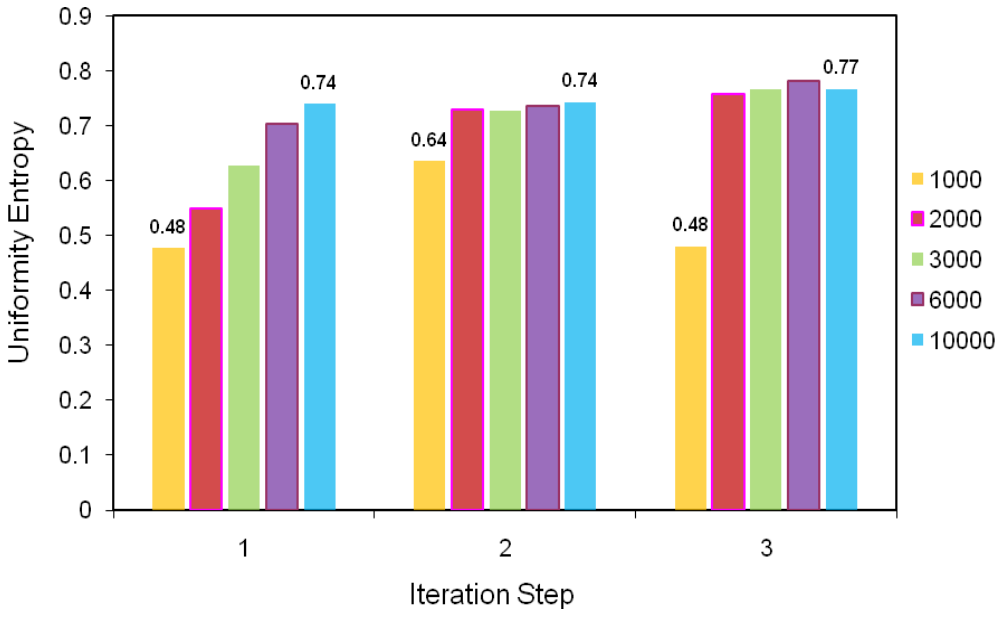


Figure 4.2(a) Uniformity of entropy in pilot sampling - Effect of N_2 (Example 1)

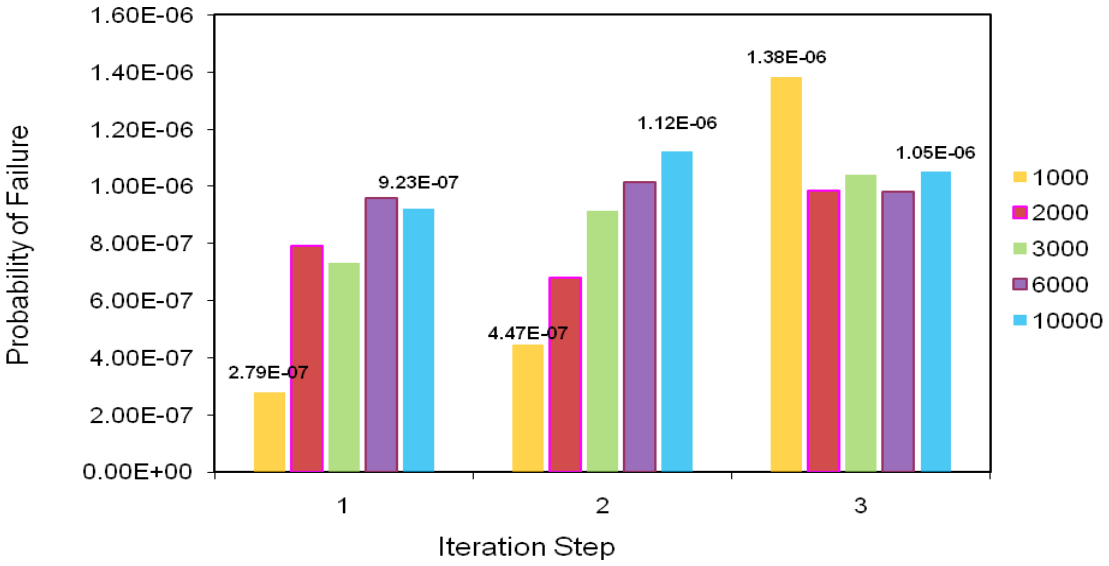


Figure 4.2(b) Probability of failure in pilot sampling - Effect of N_2 (Example 1)

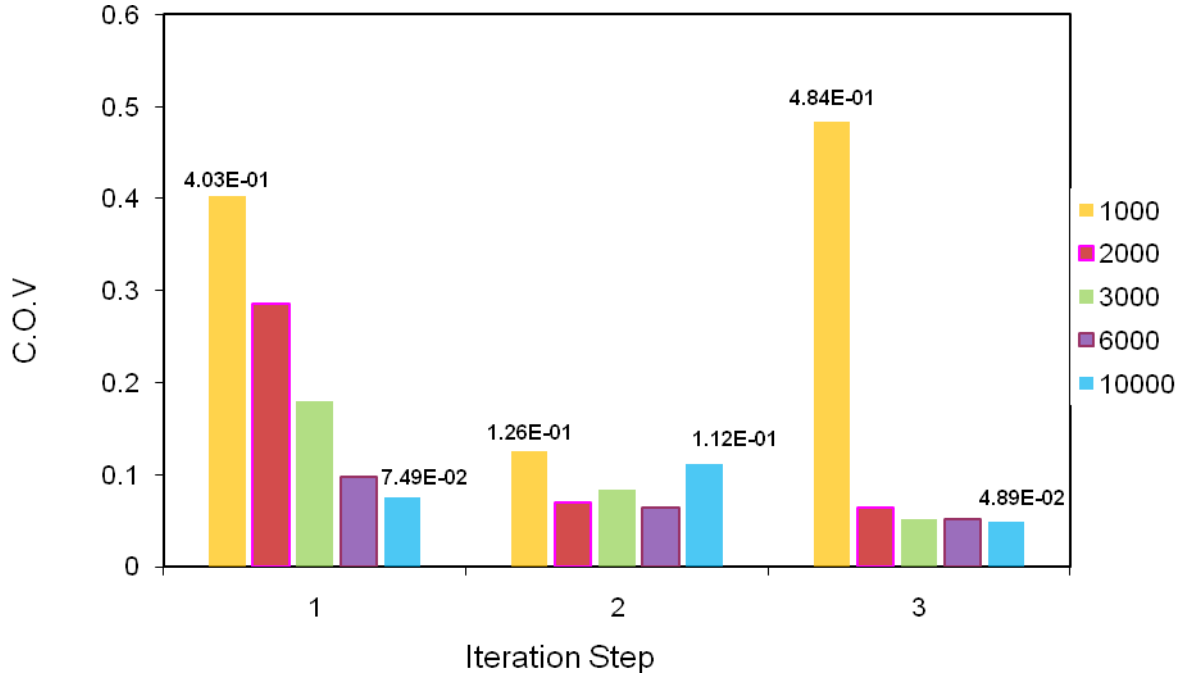


Figure 4.2(c) c.o.v in pilot sampling - Effect of N_2 (Example 1)

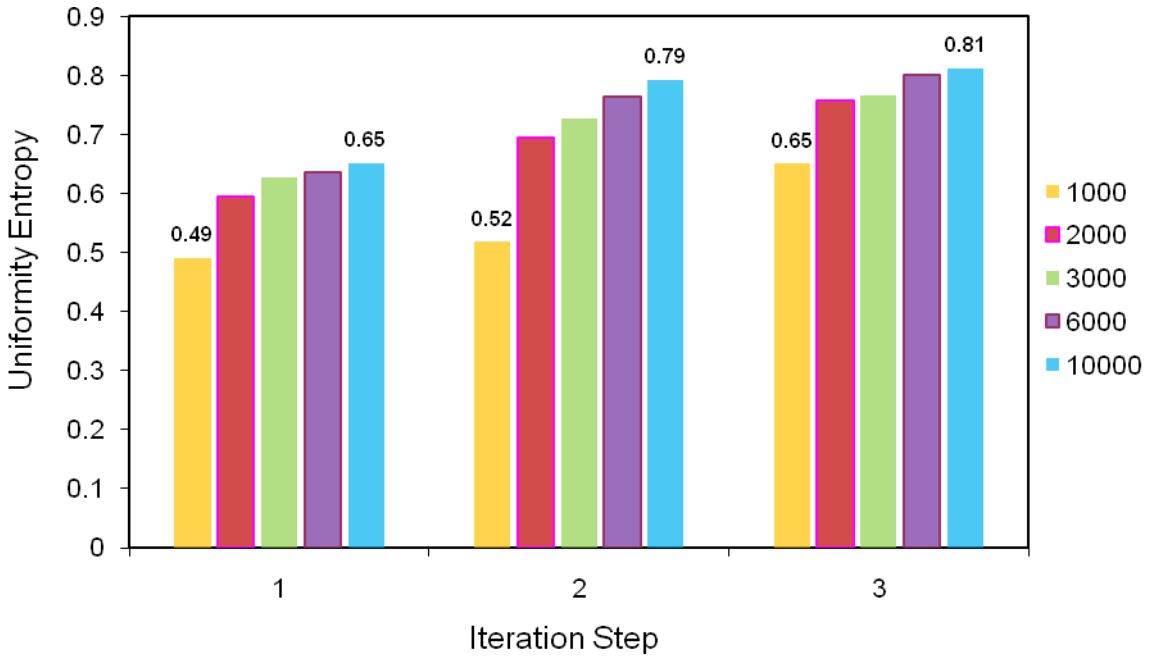


Figure 4.3(a) Uniformity of entropy in pilot sampling – Effect of *Ns* (Example 1)

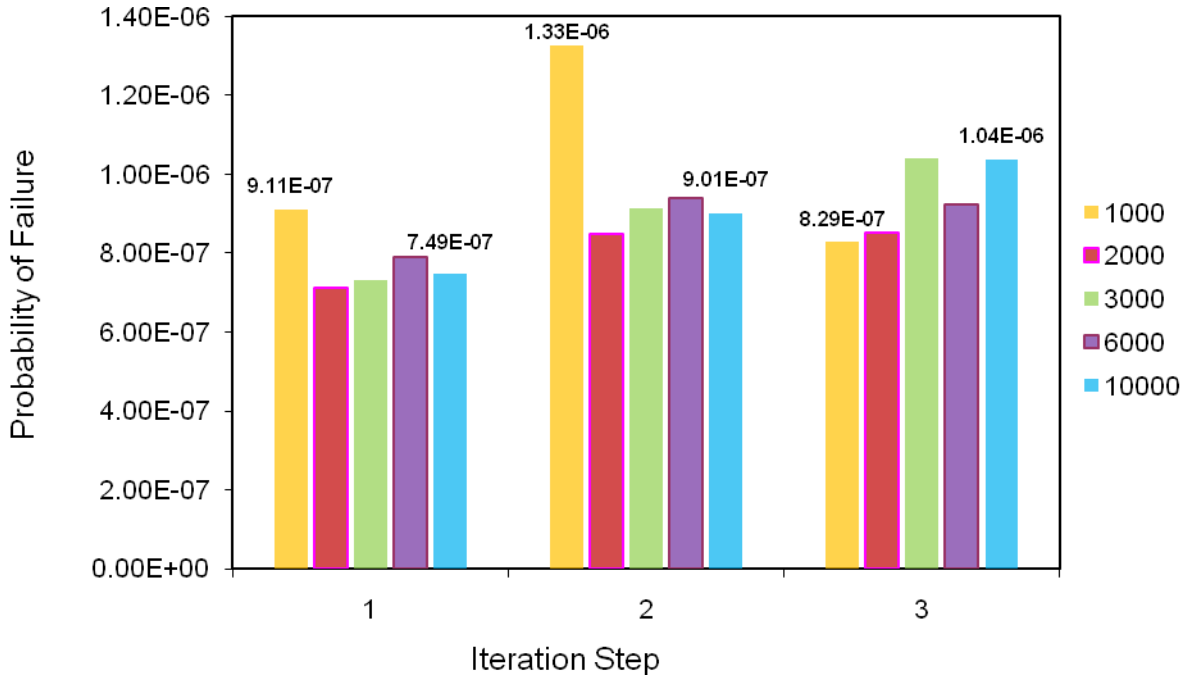


Figure 4.3(b) Probability of failure in pilot sampling - Effect of N_S (Example 1)

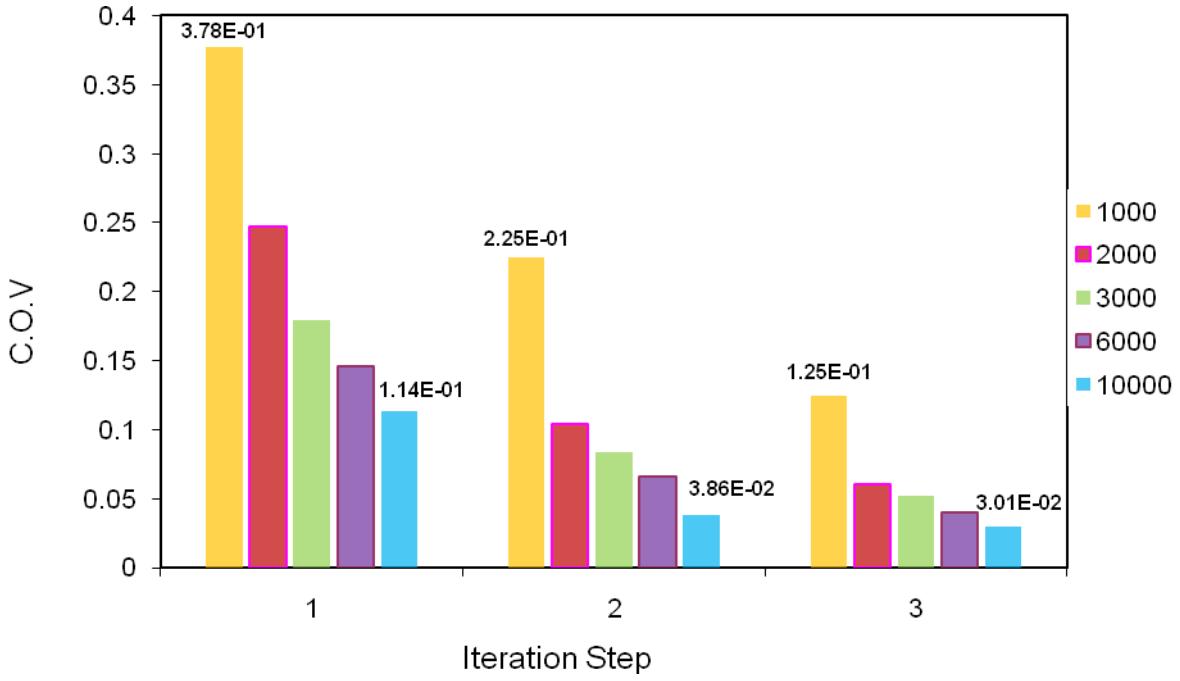


Figure 4.3(c) c.o.v in pilot sampling - Effect of N_S in (Example 1)

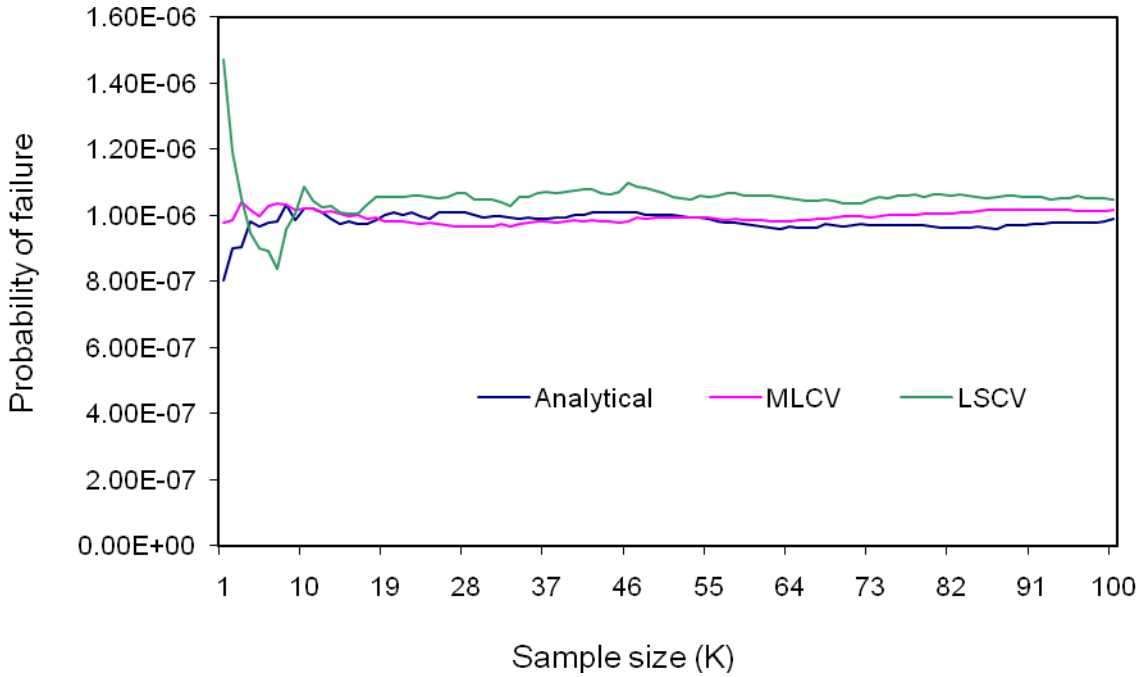


Figure 4.4(a) Probability of failure in importance re-sampling with global and local scaling (Example 1)

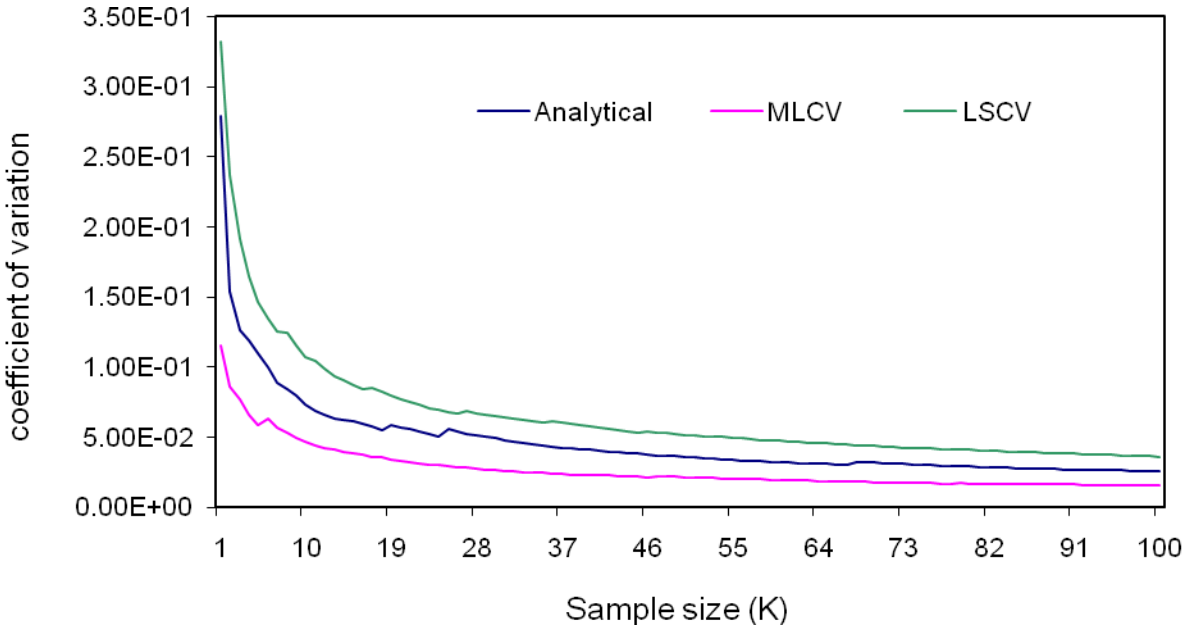


Figure 4.4(b) c.o.v in importance re-sampling with global and local scaling (Example 1)

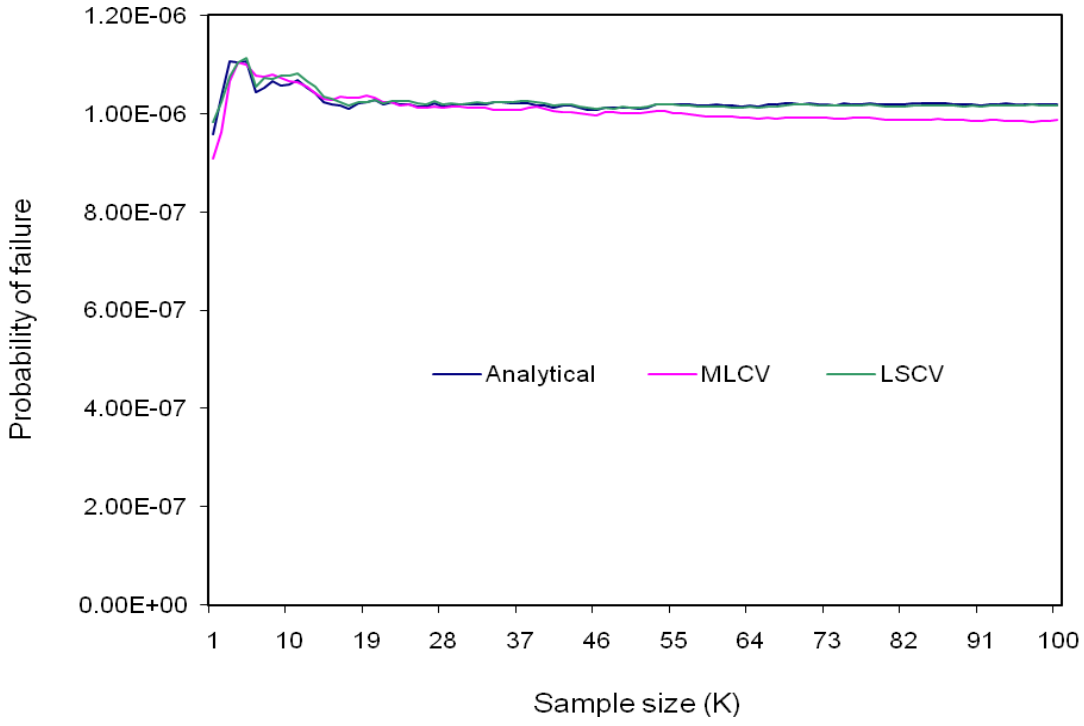


Figure 4.5(a) Probability of failure in importance re-sampling with global scaling only (Example 1)

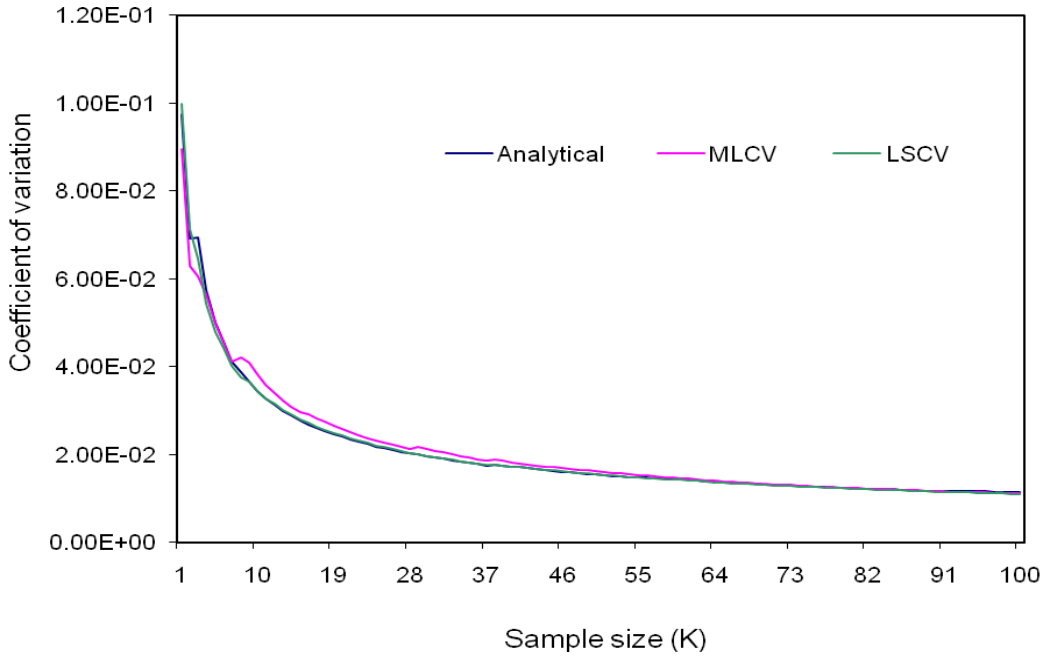


Figure 4.5(b) c.o.v in importance re-sampling with global scaling only (Example 1)

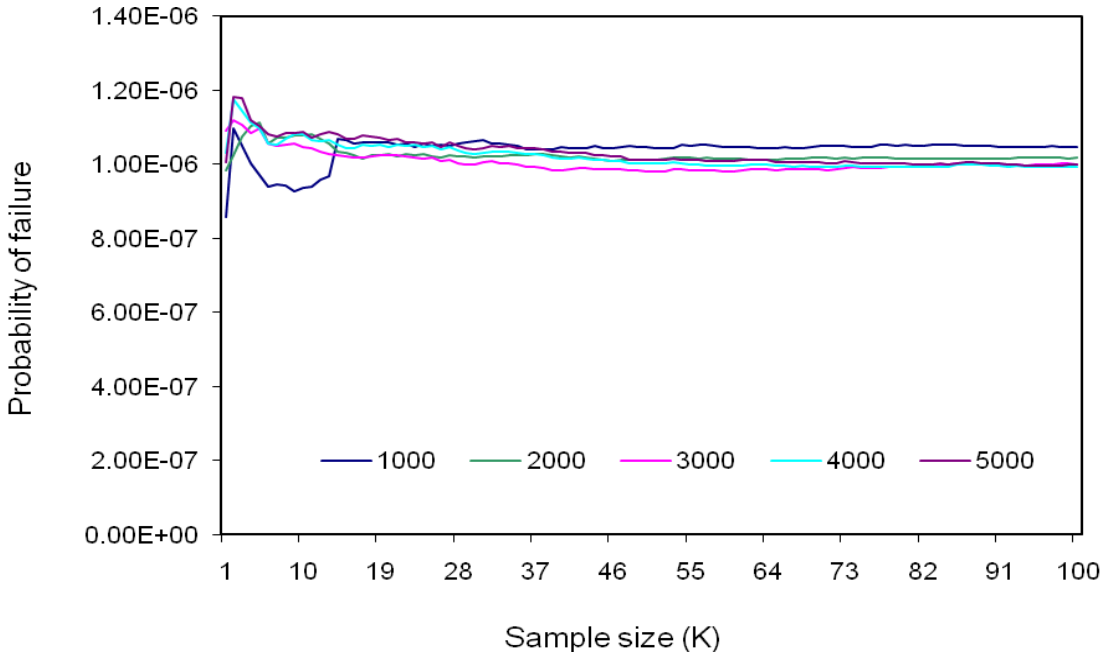


Figure 4.6(a) Probability of failure in importance re-sampling with collapsed kernels (Example 1)

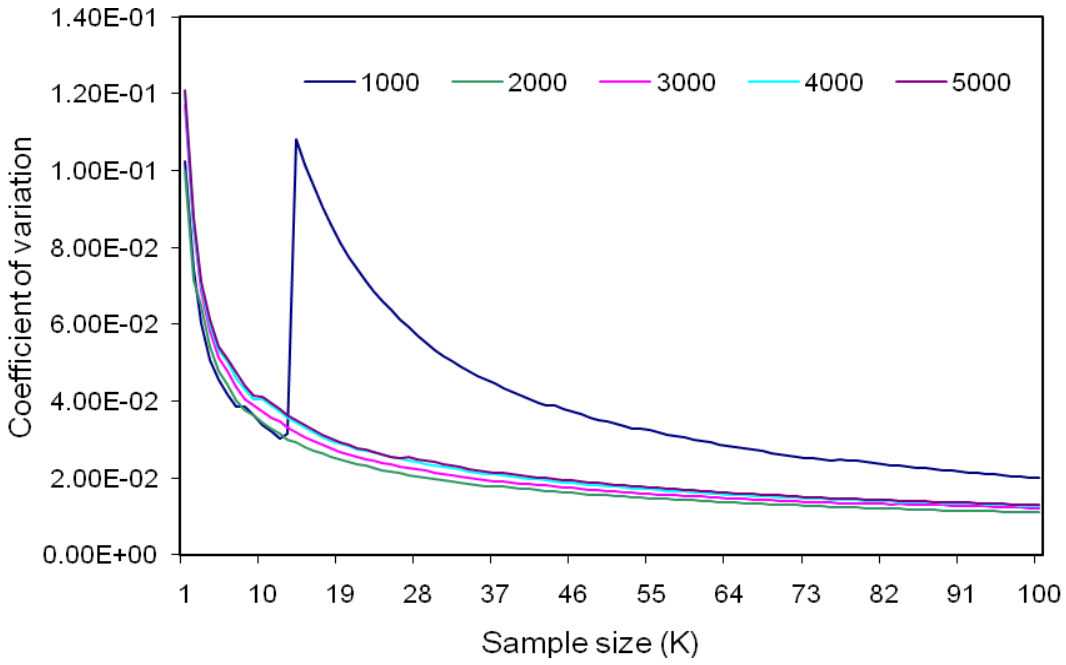


Figure 4.6(b) c.o.v in importance re-sampling with collapsed kernel (Example 1)

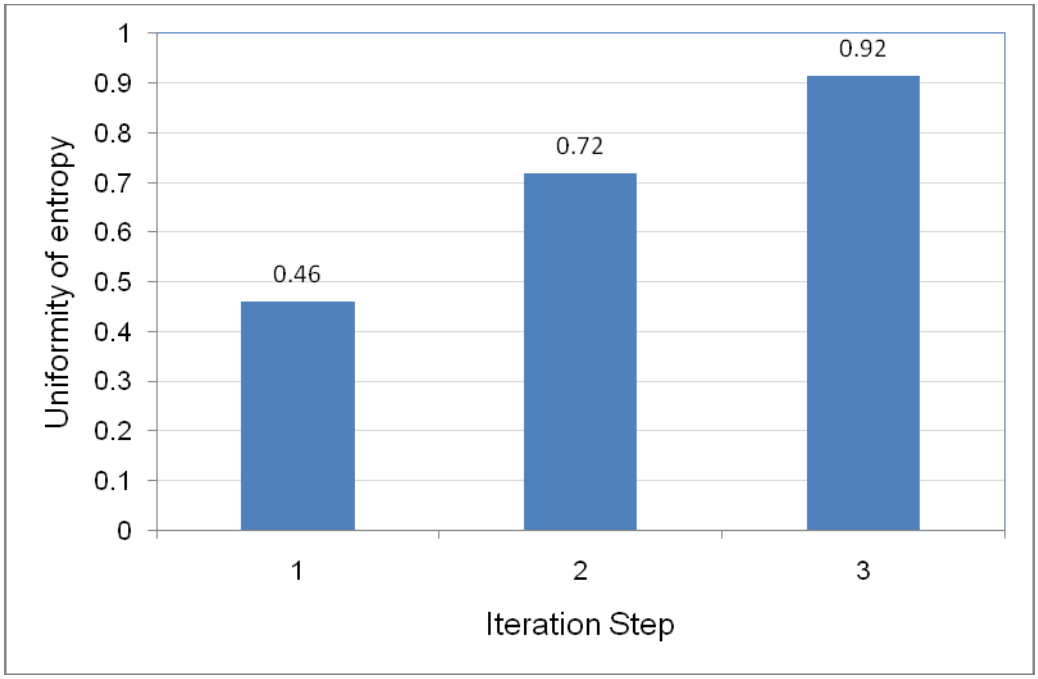


Figure 4.7(a) Uniformity of entropy in pre-sampling after convexity change (Example 1)

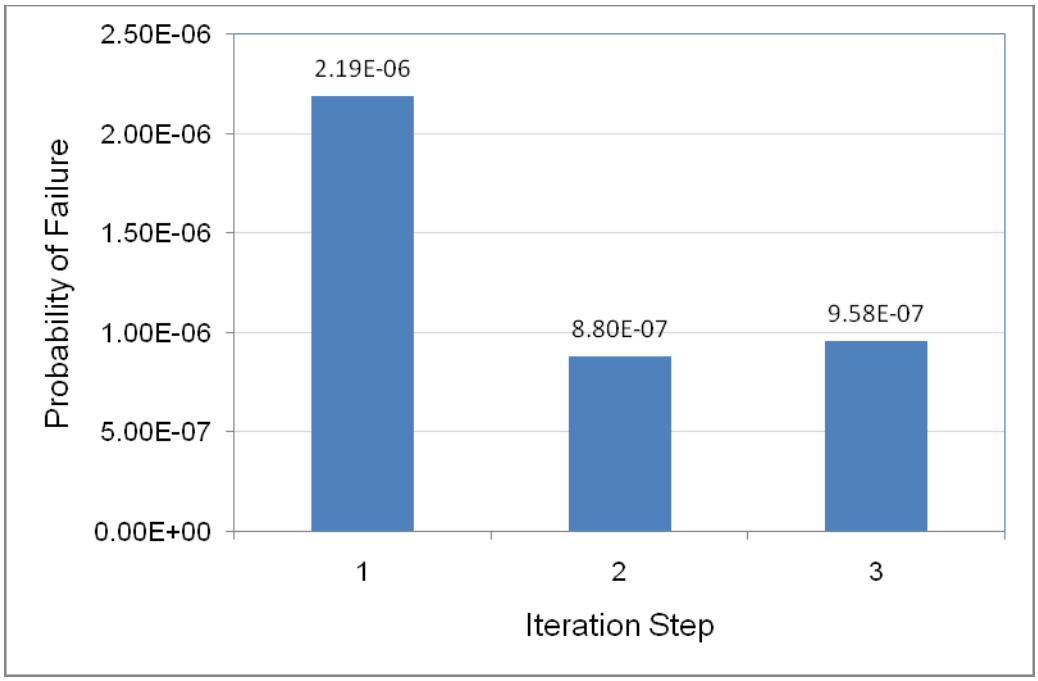


Figure 4.7(b) Probability of failure in pre-sampling after convexity change (Example 1)

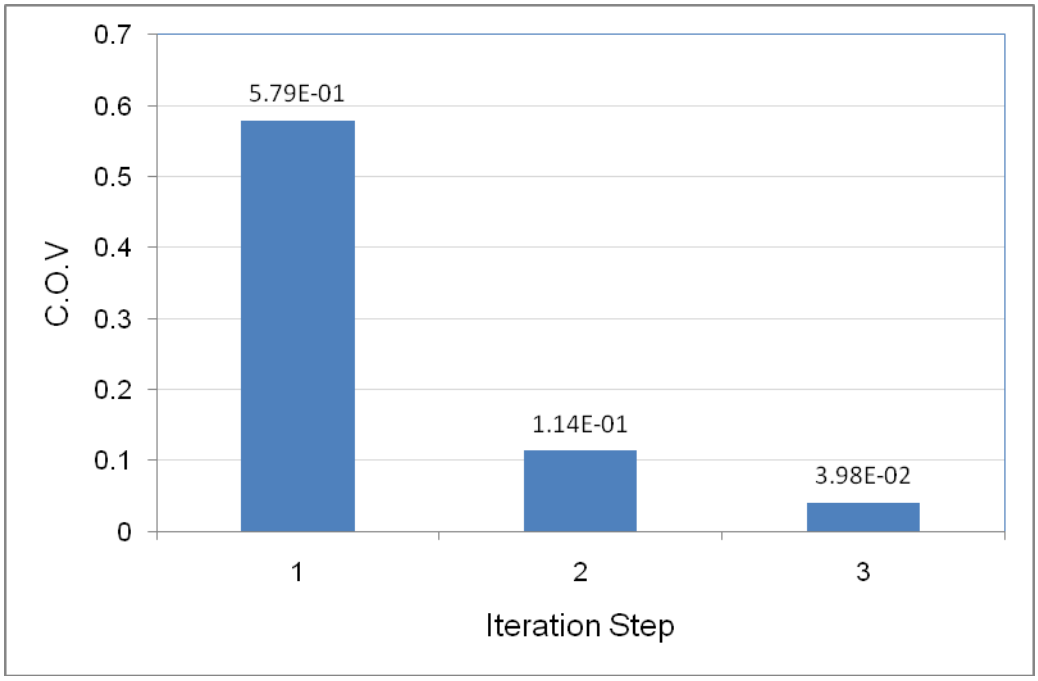


Figure 4.7(c) c.o.v in pre-sampling after convexity change (Example 1)

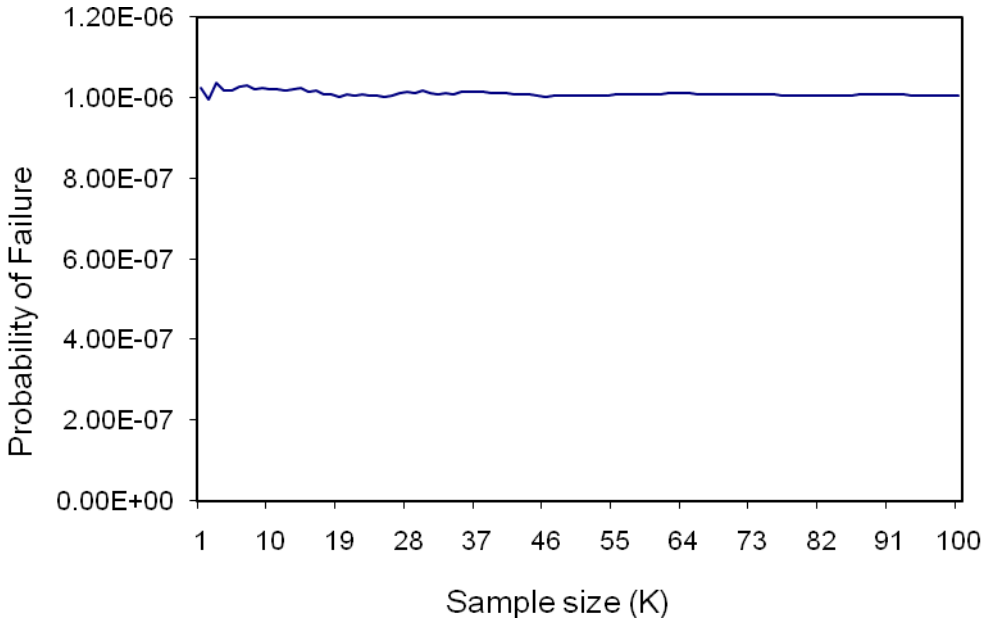


Figure 4.8(a) Probability of failure in resampling after convexity change (Example 1)

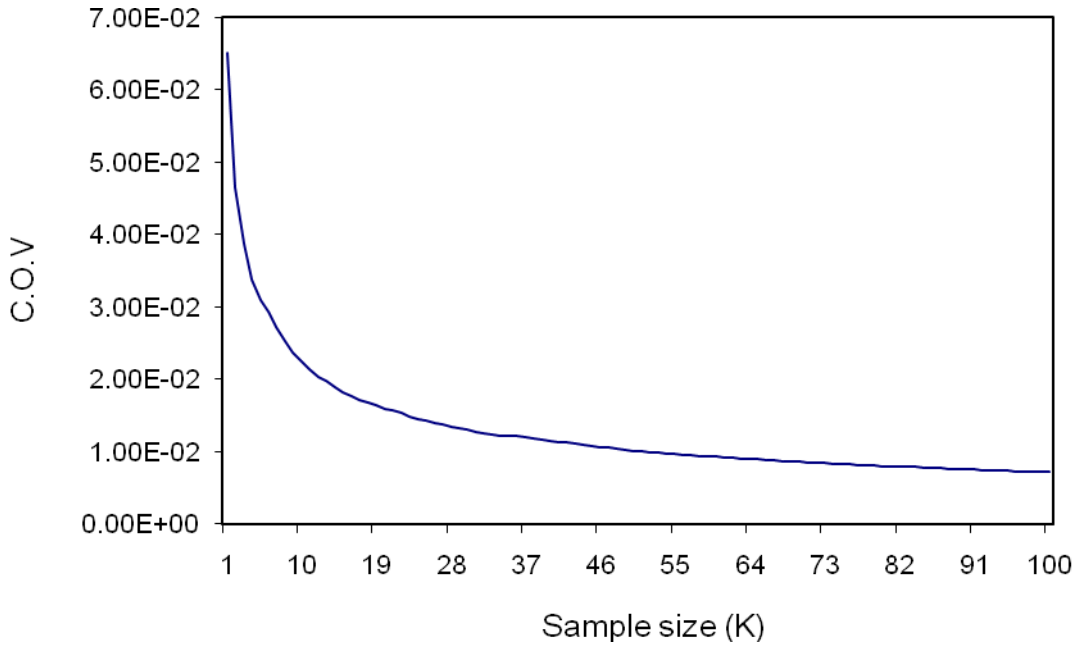


Figure 4.8(b) c.o.v in resampling after convexity change (Example 1)

4.7.2 Example 2

Consider the following limit state function

$$G(\boldsymbol{X}) = X_2 X_3 X_4 - \frac{X_5 X_3^2 X_4^2}{X_6 X_7} - X_1 = 0 \tag{4.70}$$

with correlation structure

$$\boldsymbol{C} = \begin{bmatrix} 1.0 & 0.4 & 0.3 & 0.2 & 0.1 & 0.1 & 0.1 \\ 0.4 & 1.0 & 0.4 & 0.3 & 0.2 & 0.1 & 0.1 \\ 0.3 & 0.4 & 1.0 & 0.4 & 0.3 & 0.2 & 0.1 \\ 0.2 & 0.3 & 0.4 & 1.0 & 0.4 & 0.3 & 0.2 \\ 0.1 & 0.2 & 0.3 & 0.4 & 1.0 & 0.4 & 0.3 \\ 0.1 & 0.1 & 0.2 & 0.3 & 0.4 & 1.0 & 0.4 \\ 0.1 & 0.1 & 0.1 & 0.2 & 0.3 & 0.4 & 1.0 \end{bmatrix} \tag{4.71}$$

The statistical properties of the basic variables are listed in Table 4.1

Table 4.1 Statistical properties of random variables in Example 2

Random Variable	Distribution	Mean Value	C. O. V
X_1	Lognormal	0.01	0.30
X_2	Exponential	0.3	0.05
X_3	Gumbel (Min.Type I)	360.0	0.10
X_4	Weibull (Min. Type III)	0.000226	0.05
X_5	Gumbel (Max. Type II)	0.5	0.10
X_6	Frechet (Max. Type II)	0.12	0.05
X_7	Rayleigh	40.0	0.15

This example is used to test the performance of the proposed method in problems with complicated distributions and high correlation. The result by FORM is β = 3.340, P_f = 0.419×10^{-3}. Second order analysis with direct integration by FFT gives β = 3.138, P_f = 0.849×10^{-3}. The results from the proposed kernel method are shown in Fig. 4.9 and Fig. 4.10, with different optimisation scheme. The parameters in pre-sampling are N_1 = 1000, N_2 = 1000, N_s = 1000, and N_c = 1500. In the importance re-sampling only global window width is introduced, and the number of kernel component in sampling function is collapsed from 2964 to 1500.

It can be seen that in pre-sampling the c.o.v of estimated probability of failure is a little high and it does not monotonically decrease with iteration step. To obtain better results in this stage the starting sampling size has to be increased. However the algorithm still presents a steady and fast convergence in importance re-sampling. Once again we notice that the result is not really sensitive to the optimization scheme.

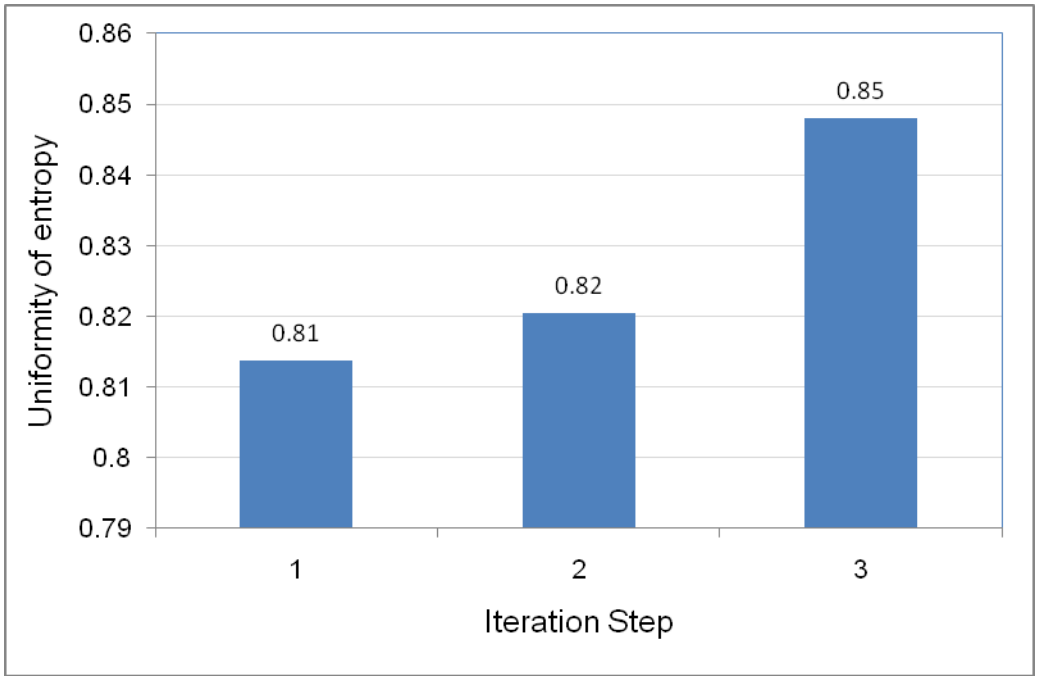


Figure 4.9(a) Uniformity of entropy in pre-sampling (Example 2)

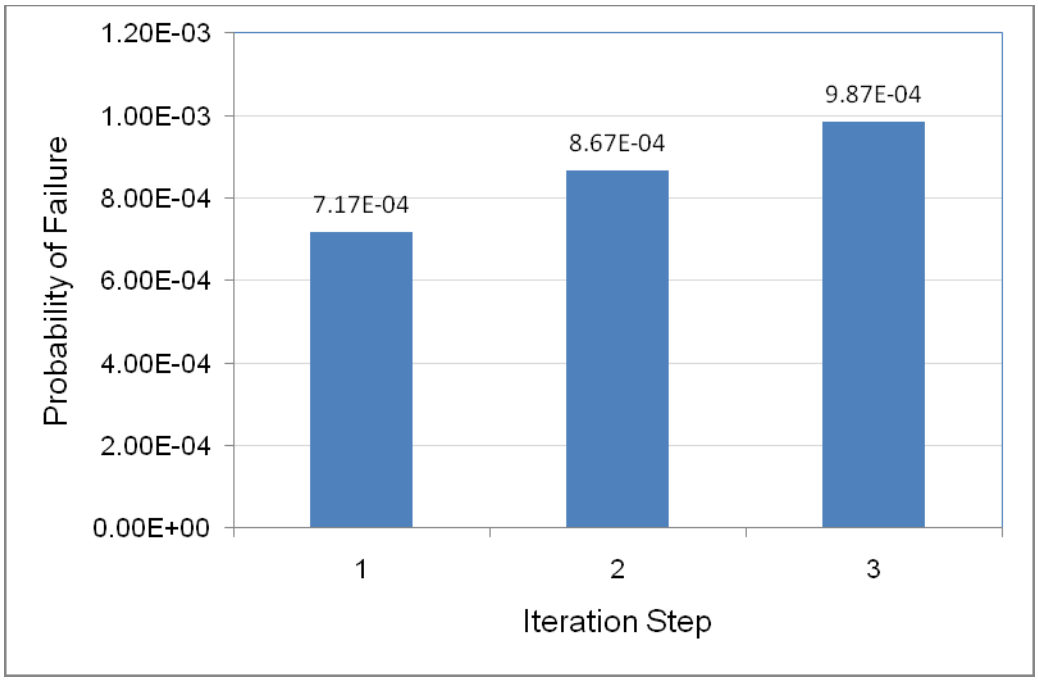


Figure 4.9(b) Probability of failure in pre-sampling (Example 2)

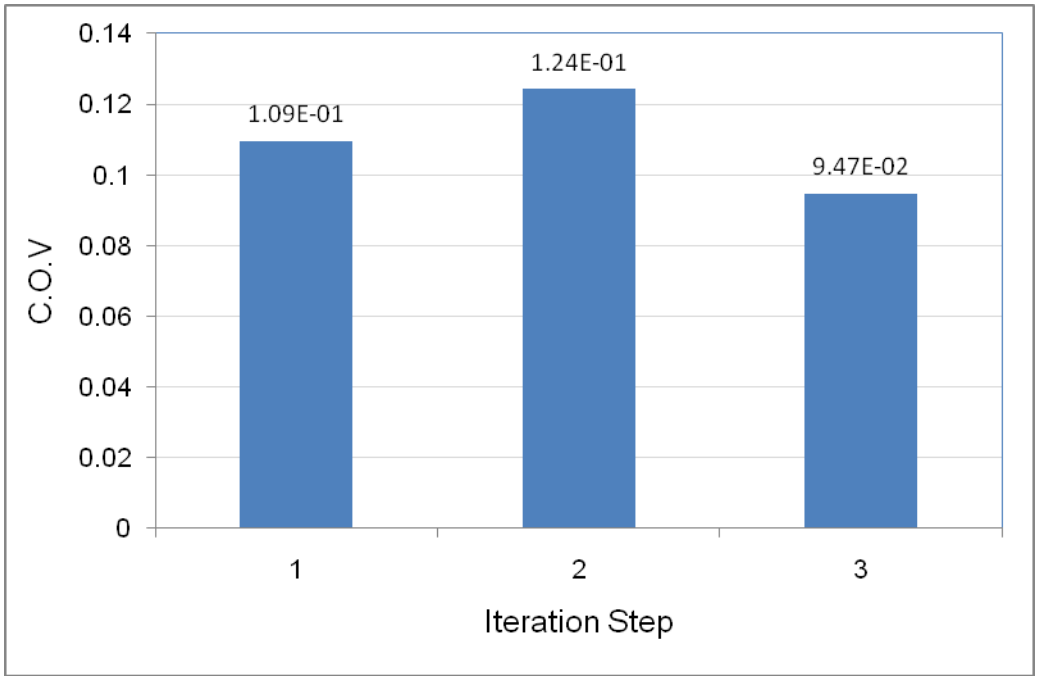


Figure 4.9(c) c.o.v in pre-sampling (Example 2)

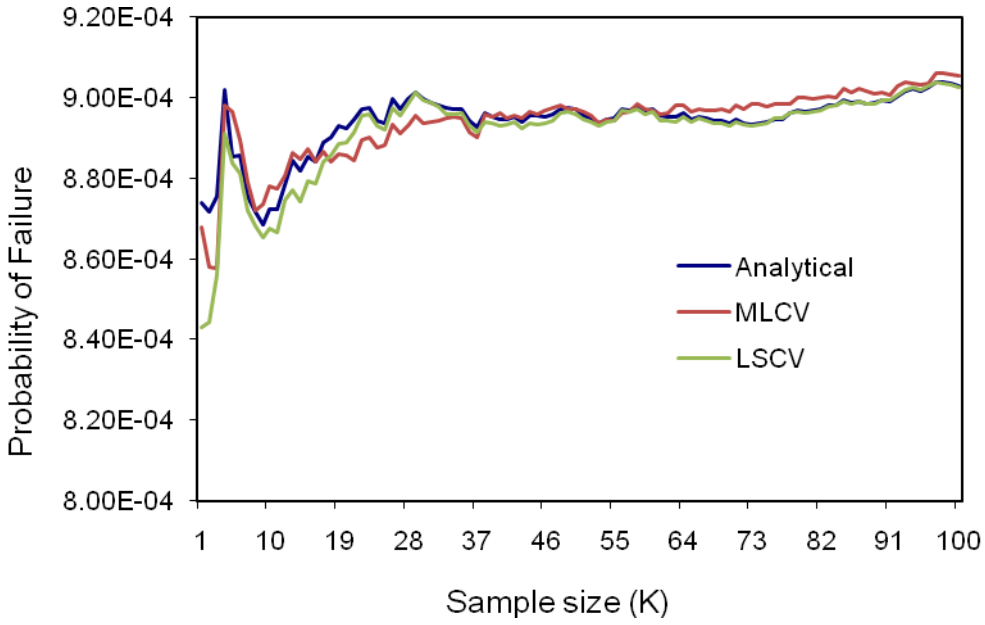


Figure 4.10(a) Probability of failure in re-sampling (Example 2)

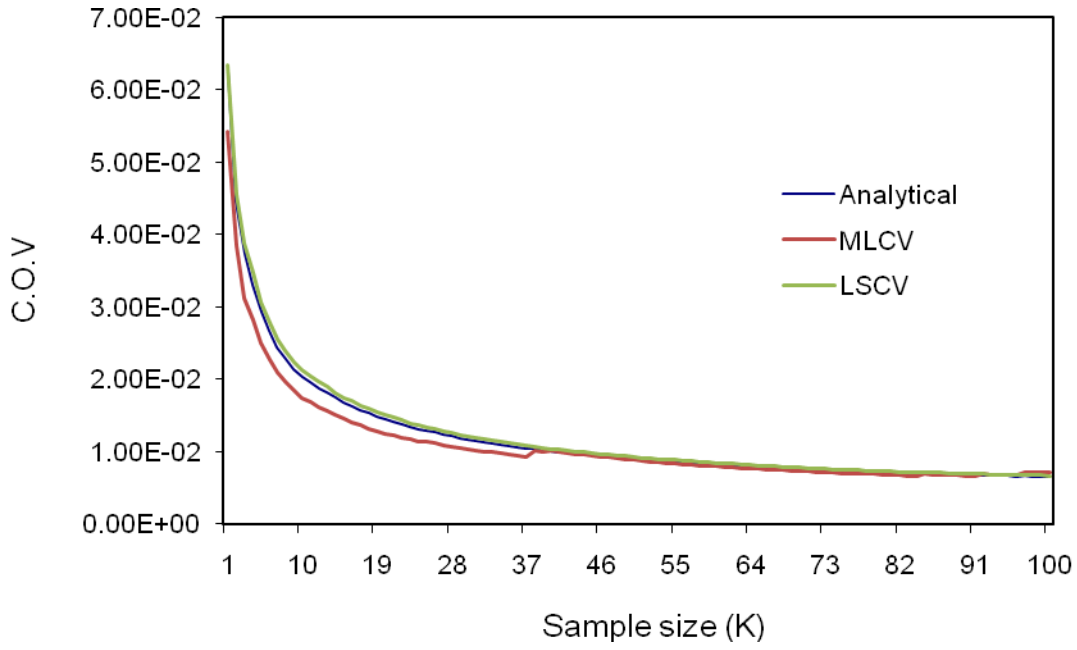


Figure 4.10(b) c.o.v in re-sampling (Example 2)

4.8 Conclusions

- the proposed importance sampling method with kernel density estimate exhibits very good precision, efficiency, and flexibility in high dimensional problem. To make the method more adaptive weighted kernel estimate is employed to approximate the most optimum sampling density. Uniformity entropy is calculated as the monitor of the goodness of fit. The accuracy of the density estimate is secured by multi-step iteration (usually 3). As is shown by the examples, using appropriate starting sampling size a good estimate of the probability of failure can be obtained in pre-sampling stage. The second part of the algorithm is optional. It is used when the first stage fails to give a satisfactory estimate. In this stage, the kernel density estimate from the first stage is collapsed and the global window width is optimised to perform importance re-sampling.

- In standard normal space local scaling factor tends to give an over smoothed density estimate which increases the variation of the estimate of the failure of probability. It should be used with special care.

- In standard normal space the global window width obtained through analytical approximation proves no much inferior to those obtained either by LSCV or by MSCV if not better. It can save us some computational time especially in pre-sampling stage.

- The accuracy of sampling density estimate is subject to the length of Markov chain, the starting sampling size, and the sampling size increment. These parameters in turn depend on the dimension of the problem and particularly the main curvatures of the failure surface in standard normal space.

References

1. Ang, G. L., Ang, H-S., & Tang, W. H. (1989). Kernel method in importance sampling density function., In A. H-S. Ang, M. Shnizuka, and G. I., Schueller, (Eds), *Proceedings of ICOSSAR 1989 Vol. II. Structural Safety and Reliability* (pp. 1193-1200). New York: ASCE.

2. Abramson, I. S. (1982). On bandwidth variation in kernel estimates-a square root law. *The Annals of Statistics*, 10(4), 1217-1223.

3. Ang, G. L., Ang, H-S., & Tang, W. H. (1992). Optimal importance-sampling density estimator. *Journal of Engineering Mechanics*, 118(6), 1146-1163.

4. Bjerager, P. (1988) Probability integration by directional simulation. *Journal of Engineering Mechanics*, 114(8), 1285-1302.

5. Bowman, A. W., & Azzalini, A. (1997). *Applied smoothing techniques for data analysis: the kernel approach with S-plus illustrations*, Oxford: Clarendon Press.

6. Breiman, L., Meisel, W., & Purcell, E. (1977). Variable kernel estimates of multivariate densities. *Technometrics*, 19(1), 135-144.

7. Breitung, K. (1991). Probability approximations by log likelihood maximization. *Journal of Engineering Mechanics*, 117(3), 457-477.

8. Bucher, C. G. (1988). Adaptive sampling – an iterative fast Monte Carlo procedure. *Structural Safety*, 5(2), 119-126.

9. Chib, S., & Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49(4), 327-335.

10. Deak, I. (1980). Three digit accurate multiple normal probabilities. *Numerische Mathematik*, 35, 369-380.

11. Ditlevsen, O., Olesen R., & Mohr, G. (1986). Solution of a class of load combination problem by directional simulation. *Structural Safety*, 4(2), 95-109.

12. Ditlevsen, O., Bjerager, P., Olesen R., & Hasofer A. M. (1988). Directional simulation in Gaussian process. *Probabilistic Engineering Mechanics*, 3(4), 207-217.

13. Ditlevsen, O., & Bjerager, P. (1989). Plastic reliability analysis by directional simulation. *Journal of Engineering Mechanics*, ASCE, 115(6), 1347-1362.

14. Ditlevsen, O., Melchers R. E., & Gluver H. (1990). General multi-dimensional probability integration by directional simulation. *Computers and Structures*, 36(2), 355-368.

15. Ditlevsen, O., & Madsen, H. O. (1996). *Structural Reliability methods*, New York: John Wiley & Sons.

16. Engelund, S., & Rackwitz, R. (1993). A benchmark study on importance sampling techniques in structural reliability. *Structural Safety*, 12(4), 255-276.

17. Fryer, M., J. (1977). A review of some non-parametric methods of density estimation. *Journal of the Institute of Mathematics and its Applications*, 20(3), 335-354.

18. Friedman, J. H., Baskett, F., & Shustek, L. J. (1975). An algorithm for finding nearest neighbors. *IEEE Transactions on Computers*, 24(10), 1000-1006.

19. Friedman, J. H., Bentley, J. L., & Finkel, R. A. (1977), An algorithm for finding best matches in logrithmic expected time. *ACM Transactions on Mathematical Software*, 3(3), 209-226.

20. Gamerman, D. (1997). *Markov chain Monte Carlo*, London: Chapman & hall.

21. Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica*, 57(6), 1317-1339.

22. Givens G. H., & Raftery A. E. (1996). Local adaptive importance sampling for multivariate densities with strong nonlinear relationship. *Journal of the American Statistical Society*, 91(433), 132-141.

23. Givens G. H. (1995). Consistency of the local kernel density estimator. *Statistics and Probability Letters*, 25(1), 55-61.

24. Hall, P., & Marron, J. S. (1987). Extent to which least-squares cross-validation minimises integrated squared error in non-parametric density estimation. *Probability Theory and Related Fields*, 74(4), 567-581.

25. Harbitz, A. (1983). Efficient and accurate probability of failure calculation by use of the importance sampling technique. *Proceedings of 4th International Conference on Application of Statistics and Probability in Soil and Structural Engineering* (pp. 825-836).

26. Harbitz, A. (1986). An efficient sampling method for probability of failure calculation. *Structural Safety*, 3(2), 109-115.

27. Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1), 97-109.

28. Hohenbichler, M., & Rackwitz, R. (1988). Improvement of second-order reliability estimates by importance sampling. *Journal of Engineering Mechanics*, 114(12), 2195-2198.

29. Ibrahim, Y. (1991). Observations of applications of importance sampling in structural reliability analysis. *Structural Safety*, 9(4), 269-281.

30. Maes, M., A. (1993). Asymptotic importance sampling. *Structural Safety*, 12(3), 167-186.

31. Melchers, R. E. (1989). Improved importance sampling for structural reliability calculation. In Ang, A. H-S., Shinozuka, M. and Schueller, G. I. (Eds.), *Proceedings of 5th International Conference on Structural Safety and Reliability* (pp. 1152-1192), New York: ASCE.

32. Melchers, R. E. (1990). Search-based importance sampling. *Structural Safety*, 9(2), 127-128.

33. Melchers, R. E. (1991) Simulation in time-invariant and time-variant reliability problems. *Proceedings of 4th IFIP WG7.5 Conference: Reliability and Optimization of Systems* (pp. 37-82), Munich.

34. Melchers, R. E. (1999). *Structural reliability analysis and prediction* (2nd ed.), New York: John Wiley & Sons.

35. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equations of state calculations by fast calculating machines. *Journal of Chemical Physics*, 21(6), 1087-1092.

36. Parzen, E. (1962). On estimation of a probability density function and a mode. *Annals of Mathematical Statistics*, 33(3), 1065-1076.

37. Park, B. U., & Marron, J. S. (1990). Comparison of data-driven bandwidth selectors. *Journal of the American Statistical Association*, 85(1), 66-72.

38. Roseblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, 27(3), 832-837.

39. Schueller, G. I., & Stix, R. (1987). A critical appraisal of methods to determine failure probabilities. *Structural Safety*, 4(4), 293-309.

40. Shooman, M. L. (1968) *Probability reliability: An engineering approach*, New York: McGraw-Hill.

41. Silverman, B. W. (1986). *Density estimation for statistics and data analysis*, London: Chapman & Hall.

42. Rubinstein, R. Y. (1981). *Simulation and the Monte Carlo method*, New York: John Wiley & Sons.

43. Shinozuka, M. (1983). Basic analysis of structural safety. *Journal of Structural Engineering*, ASCE, 109(3), 721-740.

44. Tierney, L. (1994). Markov chains for exploring posterior distributions. *The Annals of Statistics*, 22(4), 1701-1762.

45. Wand, M. P. & Jones, M. C. (1995). *Kernel smoothing*, London: Chapman & Hall.

46. Wang, G. S., Ang, A. H-S, & Lee, J. C. (1997). Adaptive Kernel method for evaluating structural system reliability. *Structural Engineering and Mechanics*, 5(2), 115-126.

47. Wertz, W., & Schneider, B. (1979). Statistical density estimation: A bibliography. *International Statistical Review*, 47, 155-175.

48. West, M. (1993). Approximating posterior distribution by mixtures. *Journal of the Royal Statistical Society* (Series B, Methodology), 55(2), 409-422.

49. Whittle, P. (1958). On the smoothing of probability density functions. *Journal of the Royal Statistical Society* (Series B, Methodology), 20(2), 334-343.

Chapter 5 Response Surface Method

5.1 Introduction

Generally speaking the strategy in reliability analysis is as follows: when the value of limit state function and its gradient can be obtained either explicitly or numerically, FORM and SORM are usually sufficient to give a reasonable estimate of the structural reliability. Unfortunately, new problem arises with the increase of scale and complexity of structural system. To obtain the gradient of failure function in such a case, design sensitivity analysis has to be carried out which is rather laborious and not supported by most commercial software products for structural analysis. On the other hand it is generally beyond any optimisation method in FORM to deal with hundreds of random variables. Therefore if higher accuracy is demanded in non-linear reliability problem or the gradient of limit state function is costly to calculate, Monte Carlo integration seems more appealing. However in reality despite the performance boosted by variance reduction techniques and adaptive iteration, the affordability of MCS is highly dependent on the scale and complexity of the problem and availability of the appropriate hardware based on cluster or grid technology.

Obviously, the above dilemma stems from the way we model the structural system and calculate responses. As a matter of fact it is not always necessary to stick to finite element method or other demanding numerical techniques all the way through reliability analysis. Rather they can be used as means to do numerical experiment that leads to a response surface function (RSF) of known form which is much easier to handle. The parameters in the assumed function can be determined after enough experimental points are obtained. The error of modelling can be reduced through adaptive iteration. This idea has blossomed into the well known response surface method (RSM). The final RSF can be subsequently used in FORM, SORM and MCS. Sensitivity results from these methods can be used to improve the design. This approach fully leverages commercial software for deterministic structural analysis.

In this chapter a review of the response surface method is given. A new response surface approach is proposed and named stepwise response surface method. It proves a quick and

reliable local approximation to the actual limit state surface in the vicinity of design point. In addition, it provides an automatic and very flexible way to pick up quadratic and cross terms according to their actual contribution. Since the sampling points are organised and used more efficiently than traditional methods, the computational effort is reduced considerably. All these advantages will be shown in the well designed examples at the end of this chapter and in the final fatigue reliability of this thesis.

5.2 Statement of problem

In some sense, RSM can be likened to a system identification method, in which a transfer function relating the input (loading and system capacity) to the output (responses in terms of displacement, stress, etc) is determined in a suitable way. Systematic numerical experiments are usually carried out in this process to either train an artificial neural network (ANN) model or to fit response surface function (RSF). A comparison study between the two approaches was done by (Gomes & Awruch 2004). However the conclusion drawn was based on a simple beam with only two random variables. In practice the use of RSF has gained far more momentum. To get a RSF with desired accuracy sufficient experiments have to be conducted with design variables in vector $\boldsymbol{X}$. At each experiment point $\boldsymbol{X}_i$ a response is observed. Although the actual response is a function of input variables, i.e., $Y = G(\boldsymbol{X})$, it is generally unavailable in closed form. The basic response surface procedure is to approximate $G(\boldsymbol{X})$ by a polynomial with undetermined coefficients. Structural analysis is performed at various points $\boldsymbol{X}_i$, in order to determine the unknown coefficients in the polynomial such that the error of approximation is minimized in the region of interest. In summary traditional RSM usually consists of the following key elements:

- Selection of RSF. Polynomial is the most popular choice because of its simplicity. An ideal polynomial should be of simple mathematical form to avoid lengthy computation in the subsequent probability analysis. Besides, the number of coefficients in the RSF should be kept reasonably low in order to reduce the number of observations required in the identification stage.

- Experimental design. The experimental points should be located in such a way that important information of the interested region can be obtained. The statistical

characteristics of random variables are usually used to locate experimental points to improve the accuracy of RSM.

- Coefficient determination. Both linear interpolation and multivariate regression can be used to determine the unknown coefficient in polynomial RSF. In fact the former is a special case of the latter where the number of experimental point is equal to the number of coefficient. To improve the accuracy the curve fitting is often coupled with adaptive iteration. At first sight, interpolation scheme seems to be more efficient since it requires less experimental data in each step of iteration. However in practice multivariate regression with redundant information can often lead to a better-positioned RSF and accelerate the convergence. It will be shown later that the computational cost in regression scheme can be reduced to the same level as in interpolation if the experimental points in one step are recycled in the following steps appropriately.

5.3 Selection of RSF

The selection of a polynomial involves the selection of its order and the terms to include. To get a well-conditioned system of linear equation for the unknown coefficients, the order of $G'(\boldsymbol{X})$ should be less than or equal to the order of $G(\boldsymbol{X})$. A higher order leads to an ill-conditioned system of equations and erratic behaviour in the sub-domains not covered by the experiment. Gavin & Yau (2008) studied high order RSF based on Chebyshev polynomial and hinted that it could be used to simulate highly non-linear limit state with multiple design points.

However any gain from high order polynomial may come at the expense of additional computation. For most structural reliability problems, one only needs to have a good approximation to the actual limit state around the design point, or the region of the failure domain where the joint probability density of design variables is relatively large and thus contribute more to the overall failure probability. Since we know neither the actual limit state function nor the actual design point, the accuracy of the reliability estimate will depend on the accuracy of the polynomial approximation in the region around the design point. Considering the accuracy of approximation and the repeated finite element analysis, a second-order polynomial is usually used. The complete quadratic model of the second order response surface is

$$G'(\boldsymbol{X}) = A + \boldsymbol{X}^T \boldsymbol{B} + \boldsymbol{X}^T \boldsymbol{C}\boldsymbol{X} \tag{5.1}$$

where A, $\boldsymbol{B}$, and $\boldsymbol{C}$ define the coefficients of the constant term, the linear terms and quadratic terms respectively.

It is the treatment of quadratic terms in Eq. (1) that features different existing algorithms. For example, a pure linear response surface is used in the paper by Kim and Na (1997), whereby $\boldsymbol{C}$ is a zero matrix. While in the Bucher's version (1990) the cross terms are omitted and $\boldsymbol{C}$ reduces to a diagonal matrix. In this paper, the quadratic terms will be introduced according to their significance level through stepwise regression.

5.4 Experimental design

The fitting of a quadratic RSF involves either multiple interpolation or regression, in which a second order polynomial of the basic variables is fitted to the observed response. Many researches on response surface analysis focused on the experimental design for regression schemes. Interpolation can be viewed as a fully saturated regression. But Guan & Melchers (2001) showed that the saturated regression was sensitive to the distance factor h. For h is equal to 0.1, 1 and 2, the obtained RSF were immensely different.

The commonly used orthogonal experimental designs in RSM are 2^n and 3^n factorial and fractional factorial designs. The 2^n factorial points for a simple two-variable (two-factor) problem are indicated in Fig. 5.1 by *. Fractional factorials are useful when the number of variables is large and hence the number of experiments that can be conducted is less than the number of combinations in the full factorial set. Augmented 2^n factorials are obtained by adding n_c points at the centre of the design. Thus the total point will be $2^n + n_c$. The centre point is indicated by $\oplus$ in Fig. 5.1. Another class of augmented design for fitting a second order response surface is central composite design. It consists of a 2^n factorial design, augmented by $n_c > 1$ central points and 2n points placed at coordinates $\pm\alpha$ along the axes. These axial points are indicated by $\ominus$ along each axis in Fig. 5.1. Hence the design consists of $(2^n + n_c + 2n)$ points (Faravelli, 1989). It should be noted that when multiple central points are used with numerical experiments a perturbation vector is often introduced such as space variation. Other experimental designs include Simplex designs

for the first order models and Equiradial designs for second order surfaces. Khuri and Cornall (1987) discussed various experimental designs in detail. These factorial designs, though efficient, lead to, unacceptably high computational efforts with the increase in the number of variables for complex systems and may become more time consuming than simulation.

Bucher and Bourgund (1990) proposed an iterative response surface approach for reliability analysis. The RSF is fitted by interpolation. The suggested experimental points are distributed along axes of each random variable. They are mean values μ_i and $\mu_i \pm h_i\sigma_i$. Here h_i is a user-defined factor and σ_i is the standard deviation of basic variables. In the paper the same design will be used while the polynomial is determined by multivariate regression instead.

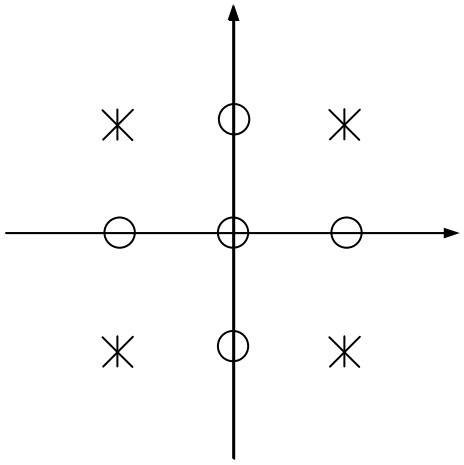


Figure 5.1 Experimental design for 2-dimentional problem

Romero et al. (2004) studied progressive lattice design which starts with n+1 sample points and gradually moves to a full multi-level factorial design. The study also features a FE interpolation approach for 2D problems. Kim and Na (1997) proposed a projection design. The sample points in each iteration are obtained by projecting the points indicated by $\ominus$ in Fig. 5.1 to the current response surface. An optimised factor f is used in their paper to take into account the effect of nonlinearity. However, this approach is subject to low speed of convergence. The idea of positing the sampling points close to the response surface was extended by Zheng and Das (2000). They introduced a small angle ε_p to

deviate the projected vector t_i, which enhances the control over sample points (see Fig.5.2).

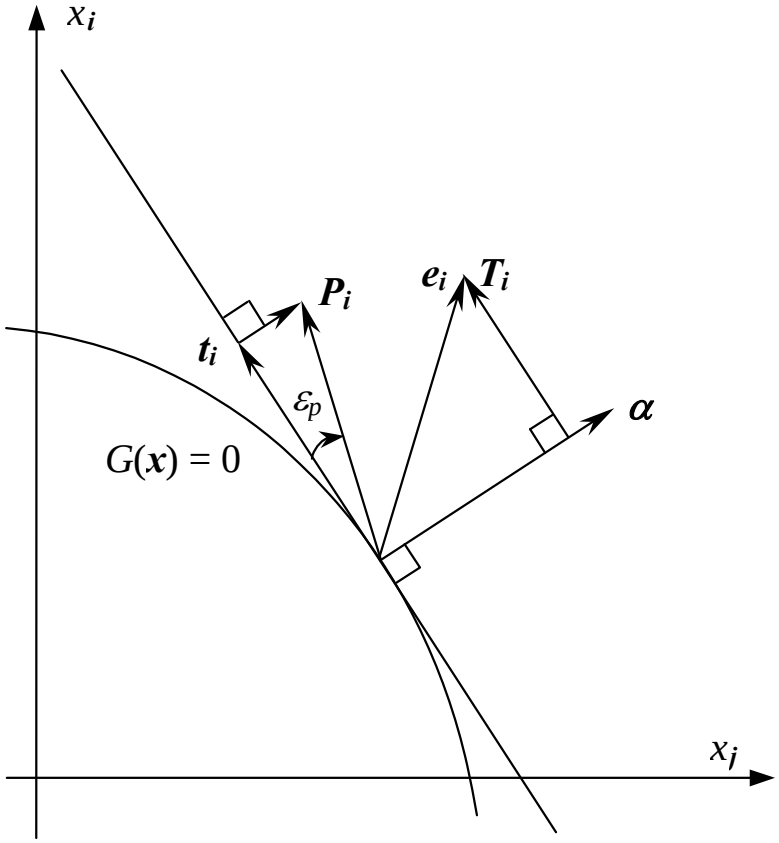


Figure 5.2 Sample points selected by vector projection

5.5 Fitting of response surface

The RSF fitted by interpolation goes through each experimental point exactly. In contrast, a redundancy is allowed for in regression. A weak form of the actual RSF is obtained in the sense of least squares. Its makes the fitted RSF less sensitive to the fluctuation in the observation. Besides, if two or more variables are collinear, the RSF from interpolation is unreliable, while this can be handled easily in regression. In the paper, stepwise regression is used. As will be shown later, it provides us with a systematic and flexible way to determine the quadratic terms in RSF.

5.5.1 Multiple regression by least squares

The regression model described by Eq. (1) is a linear one with respect to its parameters. The general form of the linear regression can be written as

$$\boldsymbol{Y} = \boldsymbol{\chi\beta} + \boldsymbol{\varepsilon} \tag{5.2}$$

where $\boldsymbol{Y}$ is a $n\times1$ vector consisting of the response variables, $\boldsymbol{\beta}$ is the coefficient factor of size $m+1$, $\boldsymbol{\chi}$ is a $n\times(m+1)$ matrix containing the experimental values of explanatory variables, and $\boldsymbol{\varepsilon}$ is the error vector whose components are composed of the lack of fit error, resulting from approximating the actual response surface by the fitted one, and a pure experimental error. We assume that $\boldsymbol{\varepsilon}$ is a zero mean, independent and homoscedastic vector, namely the standard deviations of the vector elements are constant and independent of the explanatory variables:

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \boldsymbol{I}) \tag{5.3}$$

The sum of squares of residuals is

$$S = (\boldsymbol{Y} - \boldsymbol{\chi\beta})^T (\boldsymbol{Y} - \boldsymbol{\chi\beta}) \tag{5.4}$$

By minimising S with respect to $\boldsymbol{\beta}$ we obtain the normal equation for the least square estimate

$$(\boldsymbol{\chi}^T \boldsymbol{\chi})\boldsymbol{b} = \boldsymbol{\chi}^T \boldsymbol{Y} \tag{5.5}$$

From Eq. (5) we have the unbiased estimator

$$\boldsymbol{b} = (\boldsymbol{\chi}^T \boldsymbol{\chi})^{-1} \boldsymbol{\chi}^T \boldsymbol{Y} \tag{5.6}$$

If $\boldsymbol{\chi}^T\boldsymbol{\chi}$ is not singular, the response vector from the fitted function is

$$\hat{\boldsymbol{Y}} = \boldsymbol{\chi b} \tag{5.7}$$

The vector of residuals is

$$\boldsymbol{e} = \boldsymbol{Y} - \hat{\boldsymbol{Y}} = (\boldsymbol{I} - \boldsymbol{H})\boldsymbol{Y} \tag{5.8}$$

where $\boldsymbol{H}$ is the projection matrix or the so-called hat matrix

$$\boldsymbol{H} = \boldsymbol{\chi}(\boldsymbol{\chi}^T \boldsymbol{\chi})^{-1} \boldsymbol{\chi}^T \tag{5.9}$$

5.5.2 Analysis of variance and statistical test

Vector $\boldsymbol{e}$ and $\hat{\boldsymbol{Y}}$ can be viewed as two points in the real space R^n. Their inner product is

$$(\boldsymbol{e}, \hat{\boldsymbol{Y}}) = \boldsymbol{e}^T \hat{\boldsymbol{Y}} = \boldsymbol{Y}^T \boldsymbol{H}\boldsymbol{Y} - \boldsymbol{Y}^T \boldsymbol{H}^T \boldsymbol{H}\boldsymbol{Y} = 0 \tag{5.10}$$

From Pythagoras' theory we obtain

$$\boldsymbol{Y}^T \boldsymbol{Y} = \hat{\boldsymbol{Y}}^T \hat{\boldsymbol{Y}} + \boldsymbol{e}^T \boldsymbol{e} \tag{5.11}$$

or in deviated form

$$(\boldsymbol{Y} - \overline{\boldsymbol{Y}})^T (\boldsymbol{Y} - \overline{\boldsymbol{Y}}) = (\hat{\boldsymbol{Y}} - \overline{\boldsymbol{Y}})^T (\hat{\boldsymbol{Y}} - \overline{\boldsymbol{Y}}) + \boldsymbol{e}^T \boldsymbol{e} \tag{5.12}$$

That is to say the vector of observations can be divided into two orthogonal parts consisting of the predicted values and the residuals. We can present the variation of components in a table of variance analysis (see Table 5.1).

Table 5.1 ANOVA in regression

Source of Variation	d.o.f	Sum of Squares	Mean Square
Regression	m	SSR $= \left\|\hat{\boldsymbol{Y}} - \bar{\boldsymbol{Y}}\right\|_2$	SSR/m
Residual	n-m	SSE $= \left\|\boldsymbol{e}\right\|_2$	SSE/(n-m)
Total	n	SST $= \left\|\boldsymbol{Y} - \bar{\boldsymbol{Y}}\right\|_2$	SST/m

The ANOVA table provides us the way to check how well the model fits the experimental data. One of the most important indices for this purpose is multiple correlation coefficient, which is defined by

$$R = \sqrt{\frac{U}{S_{yy}}} = \sqrt{1 - \frac{Q}{S_{yy}}} \tag{5.13}$$

where $U = SSR$, $Q = SSE$, and $S_{yy} = SST$. Another way to look at R is in terms of the angle of $\boldsymbol{Y}$ and $\hat{\boldsymbol{Y}}$, viz

$$R = COS(\boldsymbol{Y} - \bar{\boldsymbol{Y}}, \hat{\boldsymbol{Y}} - \bar{\boldsymbol{Y}}) \tag{5.14}$$

As a rule of thumb, R should be greater than 0.5 for us to have some confidence in the model. However its value depends considerably on the sample size and the number of parameters to be estimated and often over-estimates the proportion of SSR when n is small. It is hence desired that n should be at least 5 to 10 times the size of m. A more reasonable alternative of R accounting for the effect of n and m is defined as

$$F = \frac{U/m}{Q/(n-m-1)} \tag{5.15}$$

Its significance lies in that if the hypothesis $\beta_i = 0$ holds, F has a distribution of $F_{\alpha(m,\, n\text{-}m\text{-}1)}$. Obviously, the relation of these two indices is

$$R = \sqrt{\frac{mF}{(n-m-1)+mF}} \tag{5.16}$$

Likewise, the contribution of the *i*th explanatory variable can be evaluated by the sum of partial squares

$$V = U - U' = Q' - Q \tag{5.17}$$

where U' and Q' are the sum of squares obtained by eliminating x_i in the regression equation. The partial correlation coefficient is defined by

$$R_i = \sqrt{\frac{V_i}{Q'}} = \sqrt{\frac{Q' - Q}{Q'}} = \sqrt{\frac{U - U'}{S_{yy} - U'}} = \sqrt{\frac{V_i}{Q + V_i}} \tag{5.18}$$

Again under the hypothesis $\beta_i = 0$ we have the statistics

$$F_i = \frac{V_i / 1}{Q/(n - m - 1)} \tag{5.19}$$

which have a distribution of $F_{\alpha(1,\, n\text{-}m\text{-}1)}$. Apparently

$$R_i = \sqrt{\frac{F_i}{(n - m - 1) + F_i}} \tag{5.20}$$

Usually the square roots of F_i are used which have *t* distribution with the degree of freedom being *n*-*m*-1.

5.5.3 Stepwise regression

In many cases we need to discriminate from the set of possible explanatory variables a significant subset to make most of the limited experimental data. As can be imagined the number of the sub-model is massive when the original set is large. A heavily used method to limit possible models to a reasonably small number is stepwise regression, in which a path through the possible models is determined by studying one subset and then iteratively changing its participant terms. In stepwise regression a step may involve acceptance of a new candidate variable, rejection of an existing variable, or both. Gaussian elimination can

be employed to solve the normal equation and to introduce variables. The algorithm is summarised as follows

5.5.3.1 Establishment of normal equation

The centred form of normal equation Eq. (5) is

$$\begin{cases} s_{11}b_1 + s_{12}b_2 + \cdots + s_{1m}b_m = s_{1y} \\ s_{21}b_1 + s_{22}b_2 + \cdots + s_{2m}b_m = s_{2y} \\ \vdots \\ s_{m1}b_1 + s_{m2}b_2 + \cdots + s_{mm}b_m = s_{my} \end{cases} \tag{5.21}$$

where

$$\begin{aligned} s_{ij} &= s_{ji} = \sum_{k=1}^{n} (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j) \\ s_{iy} &= \sum_{k=1}^{n} (x_{ki} - \bar{x}_i)(y_k - \bar{y}) \end{aligned} \tag{5.22}$$

To obtain a more stable numerical result we can transform the normal equation to the scaled form

$$\begin{cases} r_{11}\tilde{b}_1 + r_{12}\tilde{b}_2 + \cdots + r_{1m}\tilde{b}_m = r_{1y} \\ r_{21}\tilde{b}_1 + r_{22}\tilde{b}_2 + \cdots + r_{2m}\tilde{b}_m = r_{2y} \\ \vdots \\ r_{m1}\tilde{b}_1 + r_{m2}\tilde{b}_2 + \cdots + r_{mm}\tilde{b}_m = r_{my} \end{cases} \tag{5.23}$$

where r_{ij} is the correlation coefficients between the explanatory variables or the cosines of the angle of $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ ($\boldsymbol{y}$)

$$r_{ij} = \frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}} \quad i, j = 1,2,\cdots,m,y \tag{5.24}$$

Obviously we have $r_{ii} = 1$. According to the location and scale invariance property if any of the variables is either offset or scaled by a constant, statistics R, F, and t will remain the same. The relation between the new estimates and the original ones is

$$b_i = \tilde{b}_i \sqrt{s_{yy}} \Big/ \sqrt{s_{ii}} \quad i = 1,2,\cdots,m \tag{5.25}$$

The inverse of correlation matrix $\tilde{c}_{ij}$ and the inverse of the deviation matrix c_{ij} have the following relation

$$c_{ij} = \frac{\tilde{c}_{ij}}{\sqrt{s_{ii} s_{jj}}} \tag{5.26}$$

Moreover the sums of squares are all scaled by a factor $1 / s_{yy}$.

5.5.3.2 Subset determination

Providing we have finished l steps and have introduced l variables, then the step l+1 is performed as follows

(1) Calculate the contribution of each variable by Eq. (16)

$$\tilde{V}_i^{(l)} = (r_{iy}^{(l)})^2 \Big/ r_{ii}^{(l)} = \tilde{V}_i^{(l+1)} \tag{5.27}$$

where the first "=" means the loss of contribution without $\boldsymbol{x_i}$, the second "=" means the increase of contribution if $\boldsymbol{x_i}$ is introduced.

(2) Find the variable with least value of $\tilde{V}$ in the current model and denote the value by $\tilde{V}_k^l$, then calculate the corresponding F by Eq. (19)

$$F = (n - l - 1)\tilde{V}_k^{(l)} \Big/ Q^{(l)} \tag{5.28}$$

If F is less than the critical value $F_{\alpha 2}$ under certain level of significance, x_k will be removed from the model. Otherwise, find the variable with the highest value of $\tilde{V}$ denote the value by $\tilde{V}_k^{(l+1)}$, and calculate F by

$$\begin{aligned} F &= [n-(l+1)-1]\tilde{V}_k^{l+1} \big/ \tilde{Q}^{(l+1)} \\ &= (n-l-2)\tilde{V}_k^{l+1} \big/ (\tilde{Q}^{(l)} - \tilde{V}_k^{(l+1)}) \end{aligned} \tag{5.29}$$

If F is greater than the critical value $F_{\alpha 1}$ under certain level of significance introduce x_k into the model.

(3) Make Gauss-Jordan elimination with respect to x_k either to be removed or to be introduced

$$r_{ij}^{(l+1)} = \begin{cases} r_{kj}^{(l)} \big/ r_{kk}^{(l)} & (i=k,\ j \neq k) \\ r_{ij}^{(l)} - r_{ik}^{(l)} r_{kj}^{(l)} \big/ r_{kk}^{(l)} & (i \neq k,\ j \neq k) \\ 1 \big/ r_{kk}^{(l)} & (i=k,\ j=k) \\ -r_{ik}^{(l)} \big/ r_{kk}^{(l)} & (i \neq k,\ j=k) \end{cases} \tag{5.30}$$

For the variables that have entered the model, the coefficient of regression can be calculated by

$$b_i^{(l+1)} = r_{iy}^{(l+1)} \sqrt{s_{yy}} \big/ \sqrt{s_{ii}} \tag{5.31}$$

(4) Repeat step (1) to (3). Each time the removal of variable is always considered first. Only when no more variables can be removed is the possibility of introducing new variables to be considered. In practice it is seldom that one introduced variable is removed in the following steps. It is even rarer that a removed variable is picked up later again. When the point is reached where no variable can be introduced or omitted calculate the regression constant b_0, the residual e_k, and all the other statistics to check the regression result

$$\begin{aligned} b_0 &= \bar{y} - \sum_i b_i \bar{x}_i \\ e_k &= y_k - \bar{y}_k \end{aligned} \tag{5.32}$$

In stepwise regression special attention should be paid to the following points: First the lth variable is introduced only on condition that along with the previously introduced l-1 variables it results in the least residual sum of squares. So the finally determined l

variables may not constitute the optimum subset of all the possibilities, despite the fact that it often is especially when *l* is medium.

Second if the critical value F_α is so small that all variables are included in the model, when there exists serious co-linearities between some of the variables and an ill-conditioned normal equation is formed, the stepwise regression will reject those variables of little significance automatically. This is because their contributions have been accounted for to a certain degree by the other variables in the model that are collinear with them.

Third stepwise regression can be used in a very flexible way. If we assign different weighting coefficients to V_i, the importance of the variables can be controlled. As a result some variables have priority to enter the model. On the other hand, the total number of variable can be confined to the interval $L_1<L<L_2$. So if $L<L_1$ the variable of more significance will continue to be accepted until L is equal to L_1; otherwise if $L > L_2$ the variable of less significance will continue to be rejected until L is equal to L_2. In this process F test is skipped. These tricks will be taken good advantage of in the new RSM procedure proposed later in this chapter.

5.5.4 Further comments about model selection

Besides the above-mentioned criteria, there are several other points to facilitate the model selection in RSM. It should be noted that so far we have only discussed the statistical aspects in model selection. But under no circumstances should it be used in place of prior knowledge about the relationships between the explanatory variables and the response, if any. Moreover unlike the variables in other kinds of multiple regression, those in polynomial regression have a built in structure, for example, X_i^2 and X_iX_j are related to X_i in a different way than they are related to other variables. Consequently, there are certain logical considerations, which should precede statistical ones in the choice of model. They are summarised as follows

1. A quadratic term should usually be included in the model if accompanied by the corresponding linear term. If a quadratic term stands alone in the model it implies that the turning point of the surface is at the origin with respect to this variable. This origin is determined by the design of the experiment and is to some extent arbitrary. We are not justified therefore in imposing such constrains on the fitted surface.

2. A cross-product term should usually be included in the model if accompanied by both corresponding linear terms. The presence of the cross-product term indicates that the effect on the response of changing one of the variables depends on the setting of the other and vice versa. Having accepted that the setting of both variables affects the response, we are rarely, if ever, justified in assuming that the linear components are negligible. The cross-product term is generally interpreted as accounting for variation in the response over and above the variation accounted for by the two linear terms. Its interpretation in the absence of either or both of these terms can be problematical. For example, the relationship between the response and two explanatory variables, say X_1 and X_2, including the cross-product term but omitting one linear term is

$$Y = \beta_1 X_1 + \beta_{12} X_1 X_2 + \text{function of the other variables}$$

For fixed values of X_1 and X_3, X_4, …, the response is then a linear function of X_2, with the slope coefficient proportional to X_1. Such a relationship depends on the choice of origin for X_1. As in point (1) above, we may risk putting unreal constrains to the fitted surface in this way.

However it should be noted that points (1) and (2) here do not apply to the special cases where the actual response surface precludes the linear terms. So it is always worth double checking the real contribution of the linear terms in the obtained RSF, especially when a big weighting factor is used to force their acceptance. In the actual calculation we can still start with a linear RSF as a linear approximation of the actual surface. However their contribution should drop as soon as the second terms are being introduced.

Last but not least, in the above discussion, the model selection is mainly on the basis of the residual mean square. In effect, many other criteria can be used such as C_p and PRESS (predicted residual sum of squares). We shall avail ourselves of as many criteria as possible to find a relatively optimum one to any particular problem.

5.6 Stepwise response surface method

So far, we have discussed in detail how to fit a response surface. Now let us turn to its application in reliability analysis. In fact when a response surface is obtained the second-moment method and the importance sampling method can be used to calculate the

probability of failure. As is pointed out earlier, in the conventional RSM the factorial or composite design is usually used to fit the response surface. But the computational efforts are unacceptable for complex structural systems and may become more time-consuming than direct Monte Carlo simulation. Moreover, the probabilistic characters of the original limit state function can not be properly represented by the RSF evaluated using only the sampling points distributed around the mean point.

To improve the accuracy and efficiency of RSM, Bucher (1990) suggested an alternative process of fitting the response surface. In the first step of his algorithm, the mean vector is selected as the sampling centre. The response surface thus obtained is then used to find an estimate of the design point X_D. In the following step, the new centre point $\boldsymbol{X}_M$ is chosen on a straight line from the mean values $\boldsymbol{\mu}_X$ to $\boldsymbol{X}_D$ in order that $G(\boldsymbol{X}_M) = 0$, i.e.

$$\boldsymbol{X}_M = \boldsymbol{\mu}_X + (\boldsymbol{X}_D - \boldsymbol{\mu}_X)\frac{G(\boldsymbol{\mu}_X)}{G(\boldsymbol{\mu}_X) - G(\boldsymbol{X}_D)} \tag{5.33}$$

This process is used to ensure that sampling points chosen in the vicinity of the new centre point include sufficient information of the actual failure surface. The polynomial for response surface used by Bucher is

$$G(\boldsymbol{X}) = a + \sum_{i=1}^{n} b_i x_i + \sum_{i=1}^{n} c_i x_i^2 \tag{5.34}$$

Since the number of free parameters in Eq. (34) is $2n+1$, only $2n+1$ sampling points are required to evaluate a unique response surface in each step. Once the response surface is determined, the failure probability is computed using advanced Monte Carlo simulation.

Rajashekhar and Ellinwood (1993) proposed a new idea to improve the response surface obtained from Bucher's algorithm. In their method iteration is carried out until a convergence criterion is satisfied. The distance between current centre point and the new design point is used as the criterion. They also examined the improvement of the response surface resulting from the inclusion of both cross terms and squared terms and strategies to select better sampling points. A variation of this method can be found in the studies by

Mezeau & Lemaire (1997) & Wong et al. (2005) where interpolation in Eq. (33) was done in the standard space with empirical coefficient.

Faravelli (1989) suggested a methodology to improve the polynomial response surface by taking into account a correction factor term, which represents the error between the actual function and the estimate of the response function.

Kim and Na (1997) put forth an algorithm in which the fitting points are positioned by projecting the convention sampling points on the response surface obtained in the preceding iteration. Linear response functions are utilised and the reliability levels inherent in response functions are evaluated by FORM rather than Monte Carlo simulation.

The main problem in response surface method is to reduce the computational efforts in reliability analysis while maintaining enough accuracy. For this reason, the probabilistic characteristics of the original failure surface should be described as accurately as possible, especially in the region of the most probable failure point that contributes most to the exceeding probability. Because the RSF fitted by interpolation is used in Bucher's algorithm and its improved versions, the orientation of the fitted response surface may not be well represented though the centre point is close enough to the actual failure point. In addition, when the sign of $G(\boldsymbol{\mu}_X)$ and that of $G(\boldsymbol{X}_D)$ are the same, Eq. (33) will be an extrapolation instead of interpolation. In effect, even if it remains an interpolation through the iteration the linearity assumption in Eq. (33) may be too coarse. In both cases the convergence to the actual most probable failure point can not be guaranteed. In the method proposed by Kim and Na (1997), the information included in the quadratic terms and cross terms is not taken into account, hence the convergence could be rather slow even if the failure surface can be properly oriented.

Zheng and Das (2000) proposed a new response surface approach featuring the aforementioned experimental design and the way the sample points are used. In this method, the square terms and the cross terms are introduced in a progressive manner. All the available sampling points except those generated in the very initial stage are used to obtain a well-conditioned system matrix for regression.

In the study by Gayton et al. (2003), the initial experimental point is chosen based on an understanding of the structural behaviour. The coordinates of the design points are then assumed to be random variables whose properties are statistically determined from the

database of numerical experiments. This treatment allows the design point to be located by making use of confidence intervals and defining a confidence area. Apparently failure probability calculated as such is a typical epistemic estimate.

Despite the above development in RSM, there is still something wanting. That is how to find a systematic way to choose second order terms automatically according to their actual significance instead of a trial-and-error conjecture or incomplete engineering intuition in the case of new or novelty design. In Lee and Kwak (2006) proposed a model based on sensitivity of the failure probability which was estimated by Pearson distribution system. Gupta & Manohar (2004) proposed a RSM for dealing with multi-point failure; they also proposed local importance measure to determine the importance of random variables (r.v.).

To further answer this question, a new progressive algorithm is proposed in this chapter. Stepwise regression is used to fit the response surface. In this way we can give due consideration to the square terms and the cross terms according to their respective contribution while keeping the fitted model rationally concise. Besides the "noise" resulting from the error of observation is filtered away to some degree. Since the sampling points in all but the first iterations are recycled in the following ones, the work of structural analysis is reduced considerably. The algorithm is given below.

Algorithm 5.1

1) Assign a large weighting coefficient to the contribution V of the linear terms and zero weight to the contribution of the other terms. Fit a pure linear response surface by stepwise regression as the first approximation, with the sampling points located at $\mu_i \pm f_i \sigma_i$ and their centre μ_i.

2) Calculate design point $\boldsymbol{X}^{*(1)}$ of the fitted response surface as a new sampling centre, denote the corresponding reliability index by $\beta^{(1)}$, and fit a new linear response surface by conducting experiments at sampling points $X_i^* \pm f_i \sigma_i$ and X_i^*.

3) Calculate new centre point $\boldsymbol{X}^{*(k)}$ and the corresponding reliability index $\beta^{(k)}$, k = 2, 3,…

4) Assign approximate values to $F_{\alpha 1}$ and $F_{\alpha 2}$.

5) Assign a unit weighting coefficient to the contribution of the square terms and the cross terms while keeping the priority of the linear terms. Under the given thresholds $F_{\alpha 1}$ and $F_{\alpha 2}$ the generated terms are introduced through stepwise regression, in which the sampling points of the preceding iterations as well as the newly generated ones around $\boldsymbol{X}^{*(k)}$ are used to obtain a response surface of the second order.

6) If $\left|\boldsymbol{X}^{*(k)} - \boldsymbol{X}^{*(k-1)}\right|$ or $\left|\beta^{(k)} - \beta^{(k-1)}\right|$ is less than the given tolerances, calculate the probability of failure $P_f = \Phi(-\beta^{(k)})$, otherwise go back to step 3 for next iteration.

In algorithm 5.1 the factor *f* defines the sampling range. It has a lot to do with the quality of the fitted RSF. Usually f = 3 is used in the first approximation to cover as much information in the failure region as possible. As the estimated design point approaches the actual one f can be reduced gradually by some kind of tricks until $f \approx 1$, for instance $f^{(k)} = \sqrt{f^{(k-1)}}$ (Zheng and Das 2000). It should also be noted that the information or the experimental data obtained from the second iteration is used in accumulative manner. With the iteration going on more and more experimental data will grow steadily to meet the need of fitting a robust RSF of the second order. From this point of view, it is not appropriate to introduce more terms than linear ones in the first two iterations in case the RSF obtained be a misleading one though with a high multiple correlation coefficient. If we have got enough experimental data or when the point $f \approx 1$ has been reached, but the convergence criteria are not satisfied, a hyper-sphere can be defined by $\left\|\boldsymbol{X} - \boldsymbol{X}^{*(k)}\right\| \le \rho$, within which all the experimental points will be used to fit a RSF once and for all. In practical calculation, if a very small r_{ii} in Eq. (23) is encountered the corresponding variable will not be introduced to avoid numerical difficulty in the following Gauss-Jordan elimination. The tolerance can be set as 10^{-8} considering that a zero F_α may be used.

5.7 Examples

5.7.1 Example 1

Consider the same example again in Chapter 4, section 4.7.2. It is highly non-linear in the original space and all the basic variables are correlated with each other. The tolerance used in fitting RSF is 10^{-3}. It takes four iterations to converge. The limit state function is called 60 times. The process of iteration is shown in Table 5.1. The final response surface is given in Table 5.2. The results of sensitivity analysis are shown in Fig. 5.3 against the results obtained from the actual limit state function. Here the sensitivity is expressed in elasticity factor, which is defined by

$$\frac{d\beta}{dP}\frac{u}{\beta} \tag{5.35}$$

It can be noticed that the obtained RSF approximates the actual one at the design point so well that it gives almost identical reliability index and sensitivity results as in the original problem. This, to a great extent, attributes to the square terms and cross terms, which otherwise can hardly be identified appropriately by traditional methods. We can see those variables with higher sensitivity are more liable to be included in the cross terms. This indicates that if no other methods are available to determine the cross terms, sensitivity analysis could be of great help.

Table 5.1 Intermediate result of adaptive iteration (Example 1)

Iteration Step	1	2	3	4
Sampling centre X^*	0.111381×10^{-1}	0.222728×10^{-1}	0.195076×10^{-1}	0.195110×10^{-1}
	0.287190	0.302094	0.294051	0.294069
	218.100	355.852	0.314204	314.183
	0.202408×10^{-3}	0.225328×10^{-3}	0.216280×10^{-3}	0.216308×10^{-3}
	0.448424	0.482027	0.465691	0.465677
	0.115407	0.118708	0.117216	0.117208
	36.9843	39.9793	38.8023	38.7651
$G(X^*)$	0.133517×10^{-2}	0.129709×10^{-2}	0.207387×10^{-5}	0.708467×10^{-6}
Reliability index	3.18778	3.18914	3.33974	3.33998
Failure probability	0.716846×10^{-3}	0.713475×10^{-3}	0.419283×10^{-3}	0.418930×10^{-3}

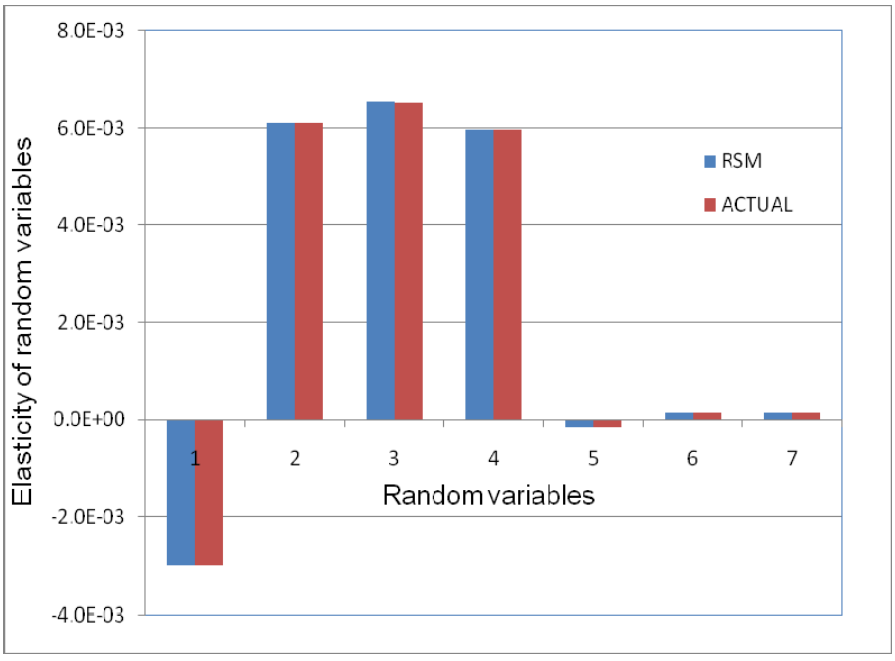


Figure 5.3(a) Elasticity of random variables (Example 1)

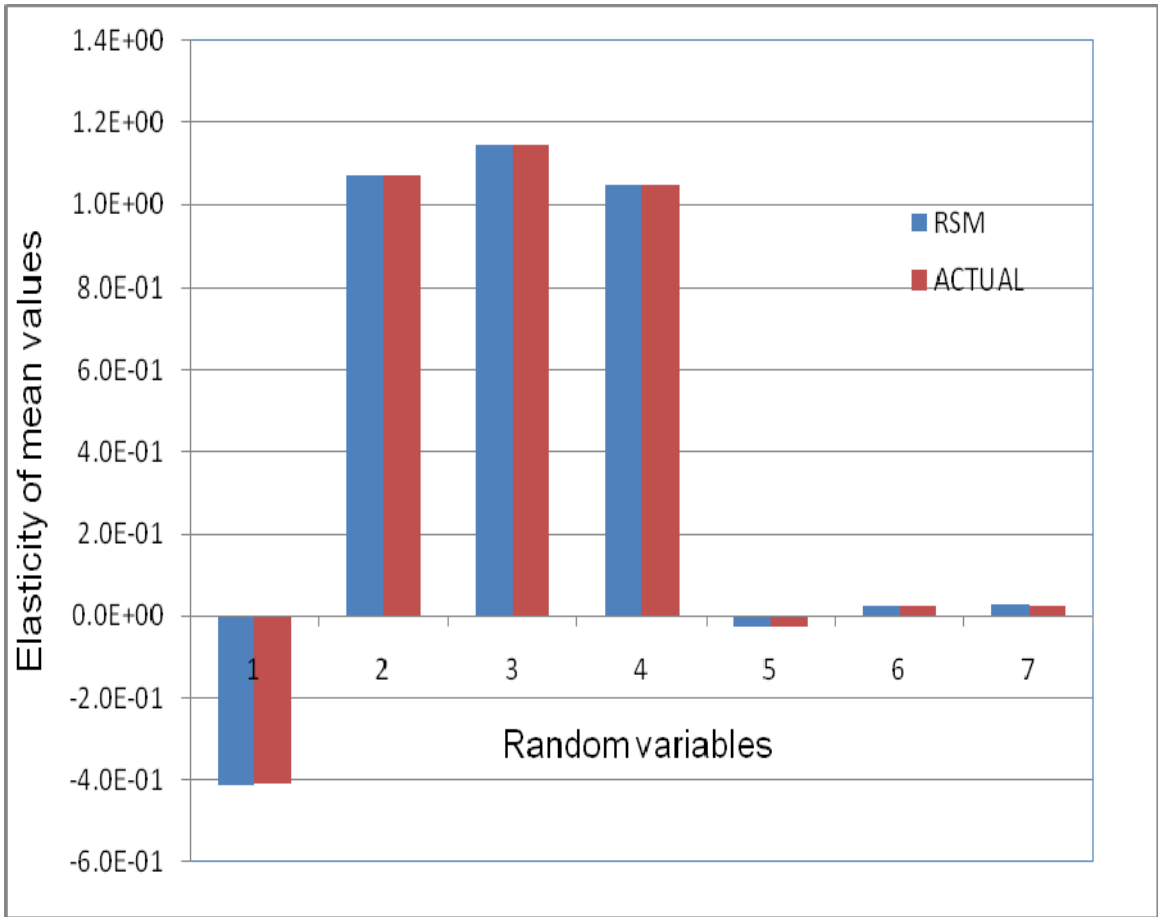


Figure 5.3(b) Elasticity of mean value (Example 1)

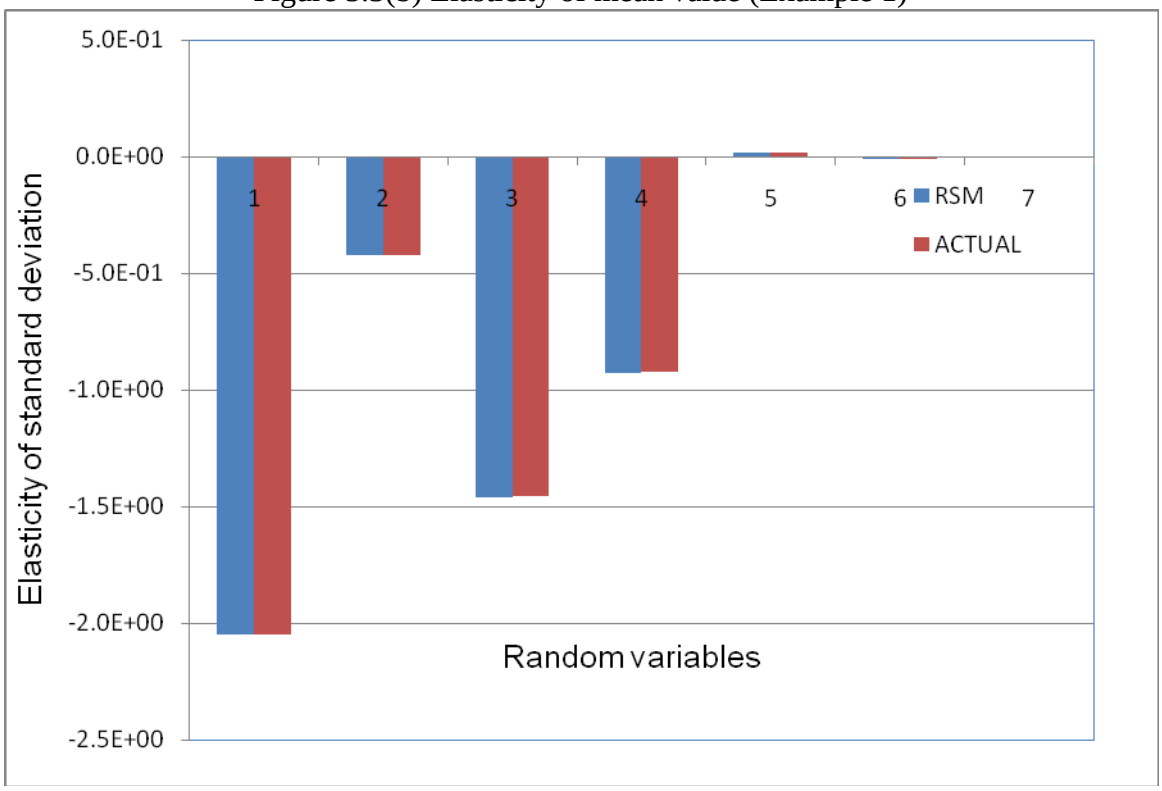


Figure 5.3(c) Elasticity of standard deviation (Example 1)

Table 5.2 Fitted Response surface (Example 1)

No.	Term	Coefficient	R
1	Constant	0.186873×10^{-1}	-
2	X_1	-0.100049×10^{1}	-0.999999
3	X_2	0.927838×10^{-1}	0.904019×10^{-1}
4	X_3	-0.433932×10^{-4}	-0.615922
5	X_4	-0.121943×10^{3}	-0.524245
6	X_5	-0.109996	-0.181396
7	X_6	0.883018×10^{-1}	0.160332
8	X_7	-0.857984×10^{-4}	-0.968689
9	X_3^2	-0.449707×10^{-8}	-0.781355
10	X_4^2	-0.680569×10^{4}	-0.183510
11	X_6^2	-0.227155×10^{-1}	-0.173019
12	X_7^2	-0.238323×10^{-6}	-0.878956
13	X_3X_2	0.131080×10^{-3}	0.494024
14	X_4X_2	0.112751×10^{4}	0.308668
15	X_4X_3	0.286642	0.963518
16	X_5X_2	0.413857×10^{-1}	0.322690
17	X_5X_3	0.314300×10^{-4}	0.267981
18	X_5X_4	-0.493039×10^{3}	-0.259187
19	X_6X_2	-0.280777×10^{1}	-0.192339
20	X_6X_3	-0.710510×10^{-4}	-0.148519
21	X_6X_5	0.165411×10^{1}	0.195933
22	X_7X_4	0.540491	0.980438
Multiple correlation coefficient: 0.999999			
Standard deviation of residual: 0.450436×10^{-5}			

5.7.2 Example 2

Consider a double bottom system shown in Fig. 5.4. It is of size 10m×10m and is stiffened by 5 girders longitudinally and 11webs transversely. The pressure of cargo is uniformly distributed. The whole structure is clamped along edge A and simply supported along edge B. The properties of random variables are listed in Table 5.3. The structure is analysed by finite element method, in which an eight-node double hull element proposed by Zheng

(1998) is used to reduce the computational time. The mesh density is 8×5. The maximum vertical displacement happens at the centre of the panel. At mean point it is 0.464cm. This agrees well with results from ABACUS (Fig. 5.5). Assume the limit state function is

$$G(\boldsymbol{X}) = 1.2\text{cm} - U_{\max} \tag{5.36}$$

where $U_{\max}$ is the maximum vertical displacement.

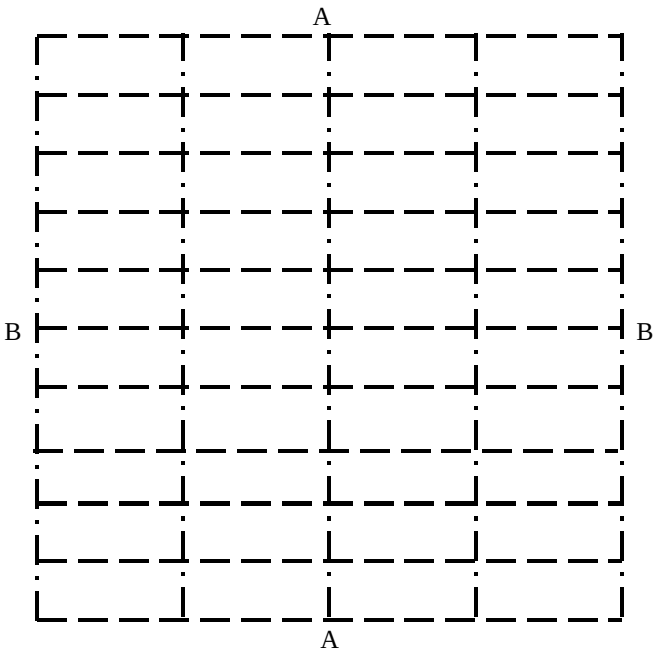


Figure 5.4 Double bottom system

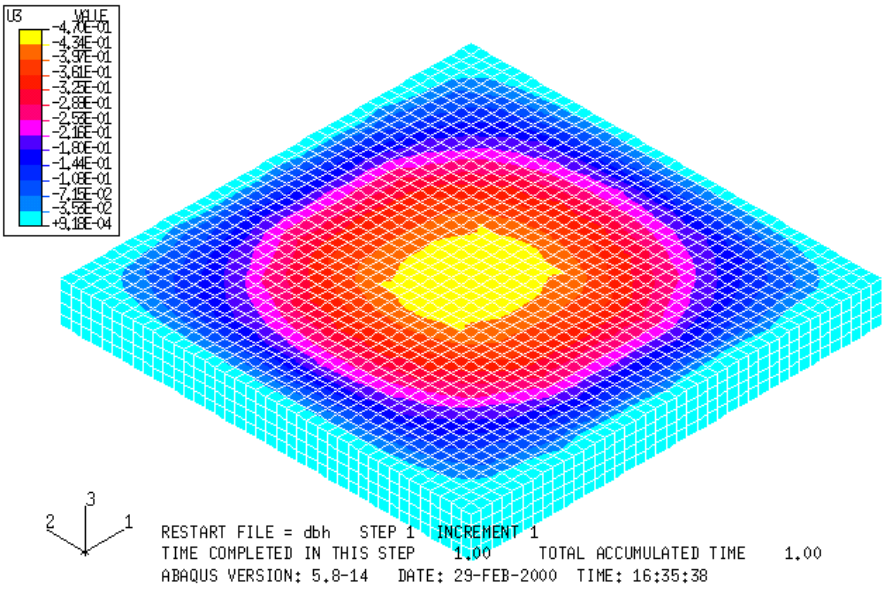


Figure 5.5 FE model in ABACUS

Given tolerance $\varepsilon = 10^{-4}$, it takes 4 step for the proposed algorithm to converge. It only costs 84 FE calculations, which cannot be beaten by even the most efficient Monte Carlo technique. The iteration process is given in Table 5.4, the final response surface is given in Table 5.5. Judging from the high multiple correlation coefficient and the low deviation of the residual, we can see the fitted response surface goes very well with the observed data. The sensitivity of reliability index with respect to random variables and distribution parameters is given in Fig 5.6. The importance of the each basic variable is quite consistent with that from the engineering intuition. The most important variables are, by order of significance, the Young's modulus, the load, the section area of the web, the thickness of the inner and outer bottom plate, and the section area of the girder. They form all the cross terms. In addition, we can see the torsional stiffness of beams plays less importance role in resisting the applied load. The torsional moment of inertia can be treated as constants.

Finally the same problem is calculated using importance sampling with kernel function as in Chapter 4. The number of samples in pre-sampling stage is given in Table 5.6. Quality and estimate of failure probability is given in Fig. 5.7. Final results in re-sampling are presented in Fig. 5.8. Comparison with STRUREL (RCP 1999), a commercial reliability software, is also given. FORM and SORM in STRUREL produces β = 3.521, P_f = 2.146×10^{-4} and β = 3.518, $P_f = 2.171\times10^{-4}$ respectively. Simple importance sampling based on SORM results gives β = 3.518, $P_f = 2.172\times10^{-4}$. The corresponding sample size is 9034, and the COV of correction factor is 0.04%. It is taken as the precise solution for comparison. It can be seen that the proposed step-wise RSM has achieved good agreements across all the methods.

Table 5.3 Statistical Properties of Double Hull System

Random variable	Distribution	Mean value	C.O.V
E	Lognormal	2.0×10^{6} kg/cm^2	0.05
T_1 (upper plate)	Normal	1.0 cm	0.05
T_2 (lower plate)	Normal	1.0 cm	0.05
I_1 (web)	Normal	8.333×10^{4} cm^4	0.05
I_2 (girder)	Normal	8.333×10^{4} cm^4	0.05
J_1 (web)	Normal	3.333×10^{1} cm^4	0.05
J_2 (girder)	Normal	3.333×10^{1} cm^4	0.05
A_1 (web)	Normal	100 cm^2	0.05
A_2 (girder)	Normal	100 cm^2	0.05
P (pressure)	Extreme I	1.0 kg/cm^2	0.25

Table 5.4 Intermediate result of adaptive iteration (Example 2)

Iteration Step	1	2	3	4
Central point X^*	2.0×10^{6}	1.97350×10^{6}	1.93792×10^{6}	1.93887×10^{6}
	1.0	0.996831	0.991902	0.992134
	1.0	0.996831	0.991902	0.992134
	8.333×10^{4}	8.33152×10^{4}	8.32921×10^{4}	8.32931×10^{4}
	8.333×10^{4}	8.32965×10^{4}	8.32441×10^{4}	8.32461×10^{4}
	33.33	33.33	33.33	33.33
	33.33	33.33	33.33	33.33
	100	99.5689	98.9035	98.9318
	100	99.9135	99.7818	99.7894
	1.0	2.56569	2.48069	2.48670
$G(X^*)$	0.735958	-0.105393×10^{-1}	0.197645×10^{-2}	-0.594933×10^{-4}
Reliability index	3.57482	3.51588	3.52157	3.52157
Failure probability	0.175234×10^{-3}	0.219147×10^{-3}	0.214496×10^{-3}	0.214501×10^{-3}

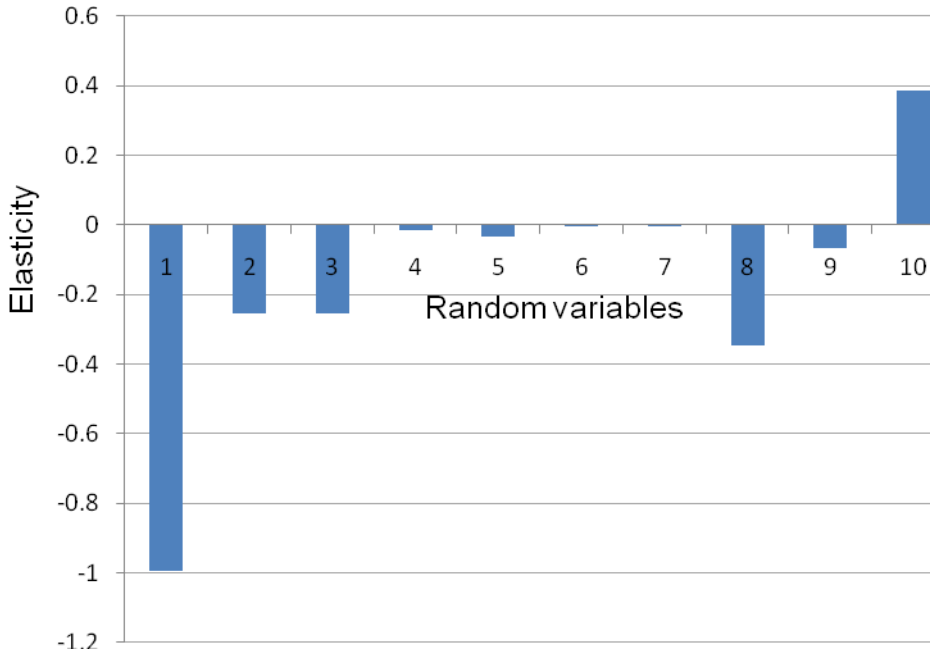


Figure 5.6(a) Elasticity of random variables

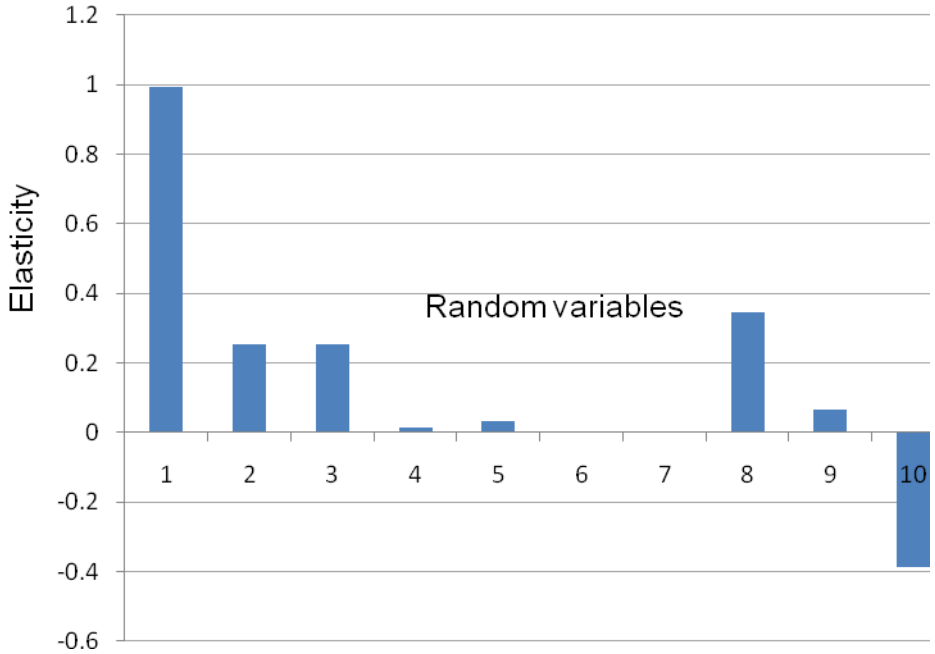


Figure 5.6(b) Elasticity of mean values

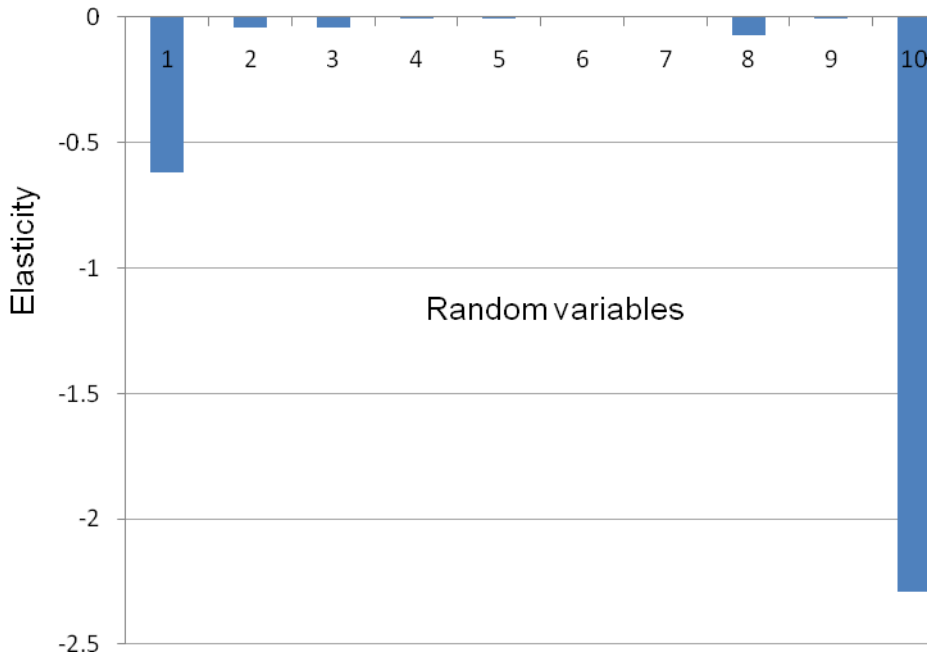


Figure 5.6(c) Elasticity of standard deviation

Table 5.5 Fitted Response Surface (Example 2)

No.	Term	Coefficient	R
1	Constant	-0.136707×10^{1}	-
2	E	0.158069×10^{-5}	0.991050
3	T_1	-0.648638×10^{-1}	-0.114979
4	T_2	-0.648638×10^{-1}	-0.114979
5	I_1	0.740521×10^{-6}	0.684108
6	I_2	0.103678×10^{-5}	0.795678
7	J_1	0.125452×10^{-2}	0.536288
8	J_2	0.125468×10^{-2}	0.536337
9	A_1	0.127595×10^{-1}	0.981342
10	A_2	0.239024×10^{-2}	0.873179
11	P	-0.126285×10^{1}	-0.996885
12	E^2	-0.320934×10^{-12}	-0.999672
13	T_1^2	-0.158169	-0.979082
14	T_2^2	-0.158169	-0.979082
15	I_1^2	-0.315600×10^{-11}	-0.554501
16	I_2^2	-0.329584×10^{-11}	-0.571186
17	J_1^2	-0.188148×10^{-4}	-0.536351
18	J_2^2	-0.188148×10^{-4}	-0.536351
19	A_1^2	-0.360031×10^{-4}	-0.995853
20	A_2^2	-0.978064×10^{-5}	-0.947844
21	P^2	-0.834210×10^{-3}	-0.535528
22	T_1T_2	0.702258	0.779035
23	EA_1	-0.387577×10^{-8}	-0.875525
24	EP	0.268908×10^{-6}	0.997910
25	A_1P	0.249040×10^{-2}	0.939094
26	A_2P	0.166057×10^{-3}	0.335779
Multiple correlation coefficient: 1.00000			
Standard deviation of residual: 0.549313×10^{-4}			

Table 5.6 Parameter configuration in pilot sampling (Example2)

Sample size in searching for PML	1000
Sample size in generating Markov chain	1500
Starting sample size in adaptive iteration	1500
Sample increment in adaptive iteration	1000
Total sample size in pilot sampling	10000

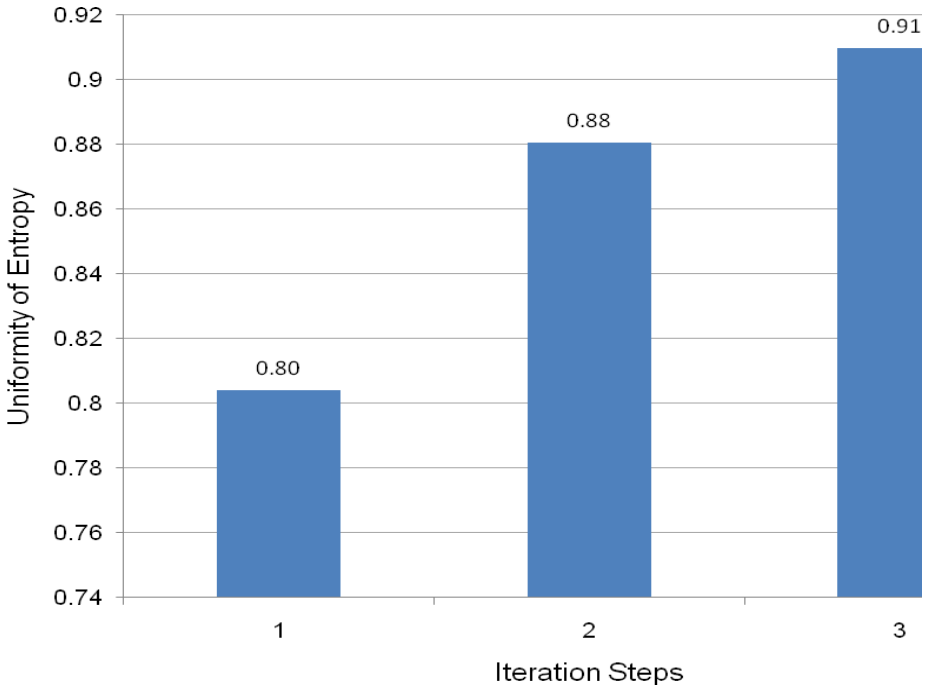


Figure 5.7(a) Uniformity of entropy in pre-sampling (Example 2)

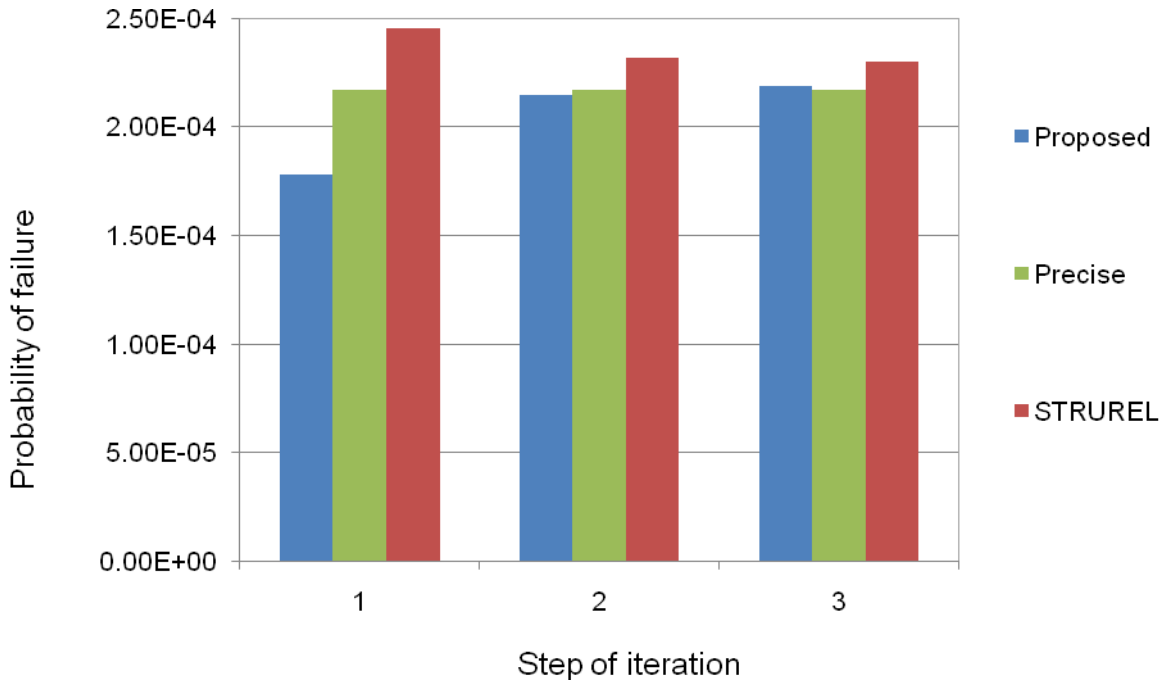


Figure 5.7(b) Probability of failure in pre-sampling (Example 2)

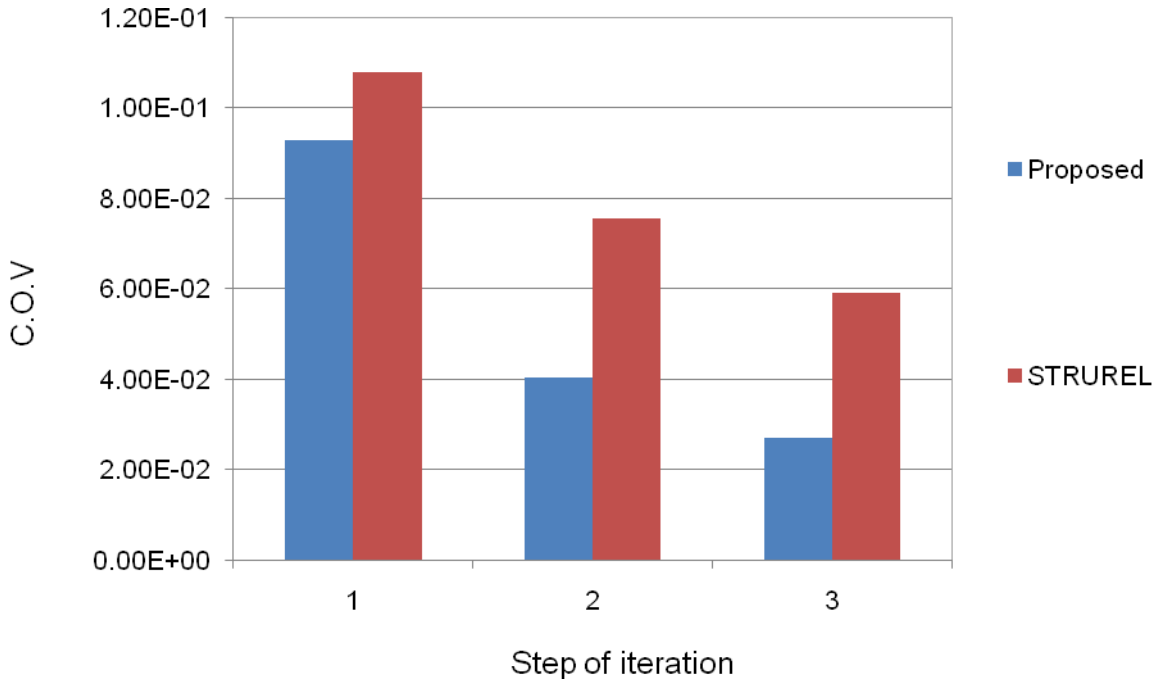


Figure 5.7(c) C.O.V in pre-sampling (Example 2)

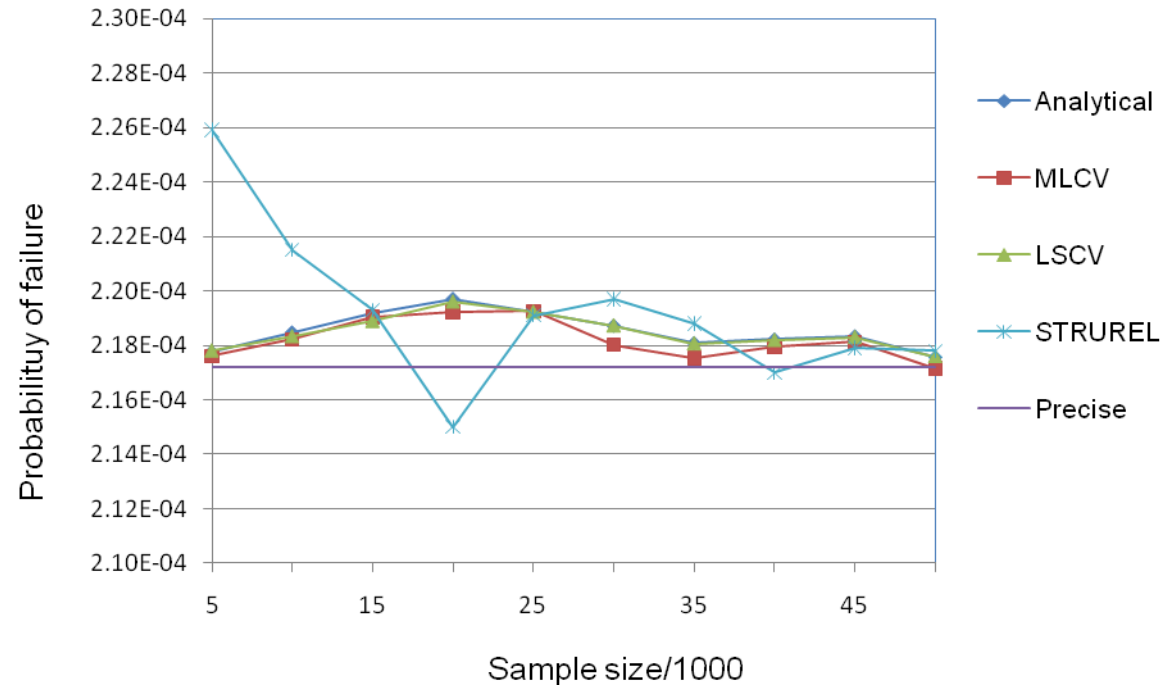


Figure 5.8(a) Probability of failure in re-sampling (Example 2)

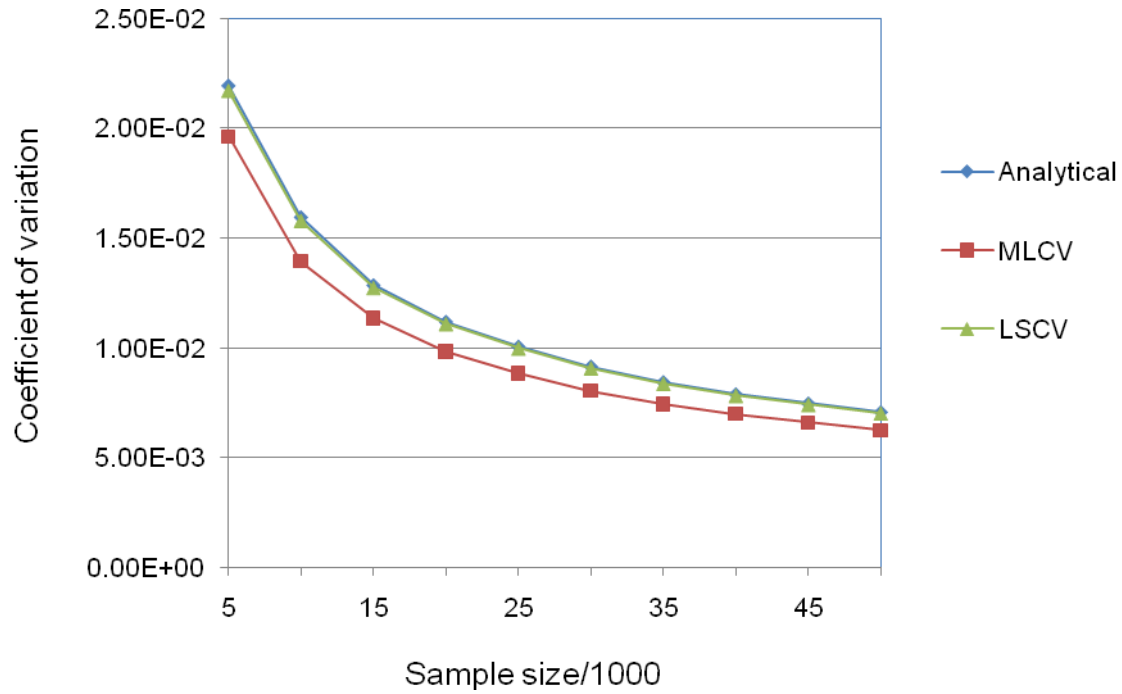


Figure 5.8(b) C.O.V in importance re-sampling (Example 2)

5.8 Conclusions

- The proposed stepwise response surface method offers a reasonable approximation to the actual limit state surface in the region of maximum likelihood. It lends us a very convenient way to take account of the quadratic terms. Since the second order information is of vital importance to pose the response surface in random space, the finally fitted RSF could be pretty close to the actual one. The sampling points in each step of iteration are recycled to reduce the work of structural analysis.

- Response surface method can work seamlessly with existing software for structural analysis. The reliability analysis of complex structural systems intractable to other methods can be performed by RSM with acceptable computational efforts. On the other hand, even those problems that can be solved by other methods, RSM can still serve as an alternative way of sensitivity analysis sometimes with unbeatable superiority.

- It should be stressed that the response surface method is by nature a modelling method, which greatly facilitates design sensitivity analysis. Besides regression, any other modelling techniques, like those from the approximation theory, can be used to improve the RSF. In this respect, some mature techniques in structural optimisation might as well be introduced into reliability analysis. This will definitely lead to dramatic variation of response surface method to suit a wide range of problems.

References

1. Bucher, C. G., & Bourgund, U. (1990) A fast and efficient response surface approach for structural reliability problems. *Structural Safety*, 7(1), 57-66.

2. Faravelli, L. (1989). Response surface approach for reliability analysis. *Journal of Engineering Mechanics*, 115(2), 2763-2781.

3. Gupta, S., & Manohar, C. S. (2004). An improved response surface method for the determination of failure probability and importance measures. *Structural Safety*, 26(2), 123-139.

4. Gavin, H. P., & Yau, S. C. (2008). High-order limit state functions in the response surface method for structural reliability analysis. *Structural Safety*, 30(2), 162-179.

5. Gayton N., Bourinet J. M., & Lemaire M. (2003). CQ2RS: a new statistical approach to the response surface method for reliability analysis. *Structural Safety*. 25(1), 99–121.

6. Gomes, H. M., & Awruch, A. M. (2004). Comparison of response surface and neural network with other methods for structural reliability analysis. *Structural Safety*, 26(1), 49-67.

7. Guan, X. L. & Melchers R. E. (2001). Effects of response surface parameter variation on structural reliability estimates. *Structural Safety*. 23(4), 429-444.

8. Kim, S. H., & Na, S. W. (1997). Response surface method using vector projected sampling points. *Structural Safety*, 19(1), 3-19.

9. Khuri, A. I., & Cornell, J. A. (1987). *Response surfaces, designs and analyses*. New York: Marchel Dekker.

10. Lee S. H., & Kwak B., M. (2006). Response surface augmented moment method for efficient reliability analysis. *Structural Safety*, 28(3), 261-272.

11. Muzeau, J. P, & Lemaire, M. (1997). Reliability analysis with implicit formulations. In C. G. Soares (Ed.). *Probabilistic methods for structural design*. Netherlands: Kluwer Academic Publishers.

12. Rajashekhar, M. R., & Ellingwood, B. R. (1993). A new look at the response surface approach for reliability analysis. *Structural Safety*, 12(3), 205-220.

13. RCP. (1999). *Strurel A Structural Reliability Analysis Program System*, Version 7.0.

14. Romero, et al. (2004). Construction of response surfaces based on progressive lattice-sampling experimental designs with application to uncertainty propagation. *Structural Safety*, 26(2), 201-219.

15. Wong, S. M. (2005). An adaptive response surface method for reliability analysis of structures with multiple loading sequences. *Structural Safety*, 27(4), 287-308.

16. Zheng, Y. (1998). Double hull finite element and its application in static and dynamic analysis of ship structures. *Journal of Dalian University of Technology*, 38(4), 419-425.

17. Zheng, Y. & Das, P. K. (2000). Improved response surface method and its application to stiffened plate reliability analysis. *Engineering Structures*, 22(5), 446-458.

Chapter 6 Fatigue Damage Model

6.1 Introduction

Fatigue models are traditionally established around SN curve approach, strain life criteria and linear elastic fracture mechanics (LEFM). SN curve is fitted through fatigue tests of pre-fabricated specimens and is represented as a two-slope linear curve in the log coordinates. So far it is most widely used in design stage not least because of its simplicity and practicability - effects of weld geometry, weld flaws and heat affected zones (HAZ) can be included in the curve statistically. Usually SN curve is obtained at mean values of logN for nominal or hot spot stresses. In real design the mean curve is shifted to the left by 2 times the standard deviation of logN so that the survivability is increased from 50% to 97.72%. Among SN curves used in ship design are UKDen (1990), IIW(1996) and proprietary ones by individual classification societies. Since the stress in ship or offshore structures is not allowed to exceed yield limit in the design codes, the SN curves currently applied in the industry can offer reliable and quick estimate in the high cycle region where $N > 10^4$. For this reason SN curve in low cycle region is often omitted in the design codes. In addition stress concentration level can be effectively controlled through careful design of the fabrication (Lloyd's Register 2004). This is often coupled by proper weld dressing such as grinding and shot penning during construction. For example design fatigue life of 17 years plus enough weld treatment at hopper knuckle is often regarded by classification societies as sufficient to meet the requirement of 25 years in the Common Structure Rules (IACS 2006).

Manson (1953) and Coffin (1954) were the first to introduce local strain criteria. It soon caused interest of marine structural engineers as they realized that though nominal stress of the ship structures is below yield stress local plastic strain can still form due to stress concentration. The difficulty of applying the approach rests on separating plastic strain range from the total strain range and the fact that real ship structures are only subject to local plastic strain when certain loading conditions are combined with encountered sea states. Because of this the damage models based on total strain range are often favoured (Tavemelli & Coffin 1962, Manson 1962). In these models the experiment data in the transition area from the low cycle to high cycle area can be well fitted. Local strain

approach is only applicable for crack initiation stage where plastic behaviour is driving force. Once the macro-crack has developed high cycle fatigue model or LEFM should be used.

Irwin (1957) revealed the relationship between stress intensity factor *K* and Griffith's rate of elastic energy release. Subsequently Paris et al. (1961) proved the paradox that fatigue cracking rate, though dominated by plastic behaviour at the crack tip, can be correlated to the elastic measure *K*. Paris law shows the predicted fatigue life is extremely sensitive to the initial conditions, including the initial flaw, the initial growth and changes in shape. As such the power law can be easily misused in fatigue life prediction if specifics of particular problem are not given due consideration (Paris 1998). This means much more will be involved in LEFM approach compared to traditional SN curve based design. That is why LEFM approach has received only guarded acceptance in marine and offshore industry and is mainly employed in fail-safe design methodology, where initial crack is known and residual fatigue life needs to be checked so that structural reliability is maintained at same level as in SN approach (DNV 2008, Cramer et al. 1992). To ease the steep learning curve faced by the industry practitioners who are used to SN approach attempts have been made to bridge the two branches. Hsu (1988) and Xu (1997) showed that LEFM can be used to construct equivalent SN curves for different initial crack length. These SN curves are then combined into a single SN equation with intercept represented as a function of the crack length. This is a refinement of similar approach recommended by IIW (1988). Petinov (2003) showed that it is possible to combine LEFM with local strain approach to produce a damage model that covers both initiation and propagation.

This chapter discusses the common damage models with emphasis on the SN approach. The effect of non-Gaussianness and bandwidth is studied through Monte Carlo simulation. A generic reliability formulation by Wirsching (1980) is also discussed.

6.2 Fatigue damage mechanism

Three stages can usually be observed in a typical fatigue test, each featuring different driving forces. In the initiation stage microscopic cracks develop first due to intensive slips in grains prone to shear deformation. Where the slips are coherent in adjacent grains the micro-cracks will trespass the threshold at a grain boundary under further cyclic loading and gradually extend to multiple grains. As the macroscopic crack initiates, the strain field

changes dramatically independent of the initial stress state of the test specimen or structure member. That is to say fatigue behaviour is dominated by the local strain field. However macroscopically the material still demonstrates strong homogeneousness even if at microscopic level numerous slips and micro-cracks have developed inhomogenously. Therefore the local strain field, when in dominance, can be identified and used to characterize the damage state in lieu of the macro-crack (Petinov 2003). As the short crack is taking form the propagating rate can increasingly be depicted by Paris law. Depending on the grain size the transition size is usually of the order of 0.1-0.125 mm. After the crack size exceeds 0.2 mm Paris law will take control (Tanaka 1987). For this reason local strain method is not suitable for welds with initial crack defects bigger than the transition size.

In stage II the macro-cracks are driven by maximum principal stress range perpendicular to the crack in the local stress field at crack tip. Crack propagation in stage II is highly dependent on the initiation condition. Once the crack outgrows the stress concentration zone at the notch and starts to affect the rigidity of the sample stage III begins. The crack propagation in stage III is almost independent of the initiation stage. As in stage II both stress and strain fields are affected by the growing crack in a way pertinent to particular test specimen. Material properties in the hysteresis loop have now relinquished control to the resistance properties of the test specimen. The transition between stage II and stage III is extremely informative. The crack size at this point can be used as criteria in fatigue design of structure details (Petinov 2003). It mainly depends on the size of stress concentration zone at the notch of the structure detail. The state of damage defined as such does not have to be compared to size that can be comfortably detected in inspection, for instance surface crack of 10-30 mm (Baker 1985) or through thickness (Huther & Henry, 1992).

6.3 Damage accumulation law

Fatigue damage is usually measured by a scalar indicator D, which is zero when the structure is intact and one at failure point. It offers a uniform criterion to quantify fatigue damage, just as the measure of probability. The general model of fatigue damage accumulation can be depicted by (Madsen et al. 1986)

$$\Delta D_n = D_n - D_{n-1} = \xi(D_1, D_2, \cdots, D_{n-1}, S_n), n = 1,2,\ldots \tag{6.1}$$

where, ΔD_n is the damage increment, and S_n is the stress range of the nth stress cycle. Here ξ is assumed to be a non-decreasing function accounting for all the fatigue contributors such as the average level of stress, load history, and environmental conditions.

If the damage accumulation is interaction-free, Eq. (1) can be rewritten as

$$\Delta D_n = \xi(D_{n-1}, S_n), n = 1,2,\ldots \tag{6.2}$$

Equation (2) shows that the damage increment relies only on the initial state of fatigue damage and the stress cycle itself. For constant S, we can further express the damage as

$$D = \eta(\frac{n}{N(S)}, S) \tag{6.3}$$

where $N(S)$ is the number of stress cycle at failure point. When η is independent of S and the damage varies slowly with the number of stress cycles we have the following kinetic equation

$$\frac{dn}{N(S)} = \frac{dD}{\eta'(\eta^{-1}(D))} \tag{6.4}$$

Integrating with both sides of Eq. (4) with substitution $D = \eta(x)$ yields

$$\sum_i \frac{n_i}{N(S_i)} = 1 \tag{6.5}$$

where n_i is the number of stress cycles with stress range S_i. This is the famous linear damage accumulation theory proposed first by Miner (1945) and Palmgren (1924) with assumption that η is linear with respect to cycle ratio n / N. In fact as is shown in Eq. (4) this is not necessary as long as η is stress-independent.

Different assumption in Eq. (1) will lead to different damage accumulation rules. A comprehensive overview of available damage accumulation models has been given by Fatemi et al (1998). However, no evidence shows that any of them is superior to others all

the time. That is why Miner- Palmgren's rule is still adopted in most rules and procedures in marine industry.

6.4 Fatigue damage based on S-N curve

In SN approach the number of stress cycles to failure N and stress range S are related by:

$$S^m N = C \tag{6.6}$$

The equivalent linear form is:

$$\log N = \log C - m \log S \tag{6.7}$$

where m and C are constants obtained through fatigue test. Applying the linear damage accumulation model due to Palmgren and Miner we have the fatigue damage of a structural detail subject to random loading of a single mode:

$$D = \sum_{i=1}^{k} \frac{np(S)\Delta S}{N_i} = \int_0^{\infty} \frac{Np(S)}{C/S^m} dS = \frac{N\mathrm{E}[S^m]}{C} \quad (k \to \infty) \tag{6.8}$$

where k is the number of stress range interval ΔS, n is the total number of stress cycle within ΔS, and E[] is the calculus of mathematical expectation.

For short-term narrow-banded Gaussian stress process, the stress range follows Rayleigh distribution:

$$p(S) = \frac{S}{4\sigma^2} \exp\left(-\frac{S^2}{8\sigma^2}\right) \tag{6.9}$$

Hence Eq. (7) becomes

$$D = \frac{\nu_0 T}{C} \Gamma(1 + \frac{m}{2})(2\sqrt{2m_0})^m \tag{6.10}$$

where ν_0 is the zero-crossing frequency, T is the service time, m_0 is zero spectral moment of stress response.

Equation (10) is the short term fatigue damage for a given sea state. The long term expectation of fatigue damage can be expressed by

$$D = \frac{\int\int\int\int\int\int T\nu(H_s,T_v,V,\theta,L)S^m f(S|H_s,T_v,V,\theta,L)f(H_s,T_v,V,\theta,L)dH_s dT_v dV d\theta dL dS}{C}$$
$$= \frac{T\nu_0\int\int\int\int\int\int \frac{\nu(H_s,T_v,V,\theta,L)}{\nu_0} S^m f(S|H_s,T_v,V,\theta,L)f(H_s,T_v,V,\theta,L)dH_s dT_v dV d\theta dL dS}{C}$$
$$= \frac{\int\int\int\int\int D(H_s,T_v,V,\theta,L)f(H_s,T_v,V,\theta,L)dH_s dT_v dV d\theta dL}{C} \tag{6.11}$$

where

H_s = significant wave height;

T_v = characteristic period;

V = ship speed;

θ = ship to wave angle;

L = loading condition;

$f(S|\ H_s,\ T_v,\ V,\ \theta,\ L)$ = conditional probability density function of stress range;

$\nu(\ H_s,\ T_v,\ V,\ \theta,\ L)$ = conditional zero-crossing rate;

$\nu_0(\ H_s,\ T_v,\ V,\ \theta,\ L)$ = unconditional zero-crossing rate;

$D\ (\ H_s,\ T_v,\ V,\ \theta,\ L)$ = conditional damage;

Fatigue damage at given sea state (conditional damage) $D(H_s,\ T_v,\ V,\ \theta,\ L)$ can be calculated following Eq. (10):

$$D(H_s,T_v,V,\theta,L) = \frac{T\nu(H_s,T_v,V,\theta,L)}{C}\Gamma(1+\frac{m}{2})(2\sqrt{2m_0})^m \tag{6.12}$$

where the moment of spectrum m_0 should be obtained from the spectral analysis of the corresponding sea state. The total number of stress cycle can be obtained by

$$N_s = T\nu_0 = \int\int\int\int T\nu(H_s,T_v,V,\theta,L)f(H_s,T_v,V,\theta,L)dH_s dT_v dV d\theta dL \tag{6.13}$$

It implies

$$\int\int\int\int \frac{\nu(H_s,T_v,V,\theta,L)}{\nu_0} f(H_s,T_v,V,\theta,L)dH_s dT_v dV d\theta dL = 1 \tag{6.14}$$

It can be readily seen the equivalent unconditional long term distribution of stress range is

$$\begin{aligned} p(S) &= \int\int\int\int \frac{\nu(H_S,T_V,V,\theta,L)}{\nu_0} f(S|H_S,T_V,V,\theta,L) f(H_S,T_V,V,\theta,L)dH_S dT_V dV d\theta dL \\ &= \int\int\int\int r(H_S,T_V,V,\theta,L) f(S|H_S,T_V,V,\theta,L) f(H_S,T_V,V,\theta,L)dH_S dT_V dV d\theta dL \end{aligned} \tag{6.15}$$

where r is defined as the relative peak response rate. The discretised long term distribution takes the form

$$D = \frac{\nu_0 T}{C} \Gamma(1+\frac{m}{2}) \sum_{n=1}^{\text{all load case}} g_n \sum_{i=1,j=1}^{\substack{\text{all seastates} \\ \text{all headings}}} p_{ij} r_{ijn} (2\sqrt{2m_{0ijn}})^m \tag{6.16}$$

where

g_n = the fraction of design life in the nth load case,

p_{ij} = the conditional joint probability of sea-state i and heading j,

r_{ij} = ν_{ij} / ν_0, the ratio of the stress crossing rate in element (i, j) to average crossing rate, and

m_{0ij} = zero spectral moment of stress response.

The conditional probabilities in Eq. (16) are usually calculated through voyage simulation.

The generic form of stochastic fatigue formulation based on S-N curve and Miner's law can be written as

$$D = \frac{\nu_0 T}{C} \int_0^\infty S^m F_S(S) dS \tag{6.17}$$

And the discretised form is

$$F_S = \sum_{n=1}^{\text{all load case}} g_n \sum_{i=1,j=1}^{\substack{\text{all seastates} \\ \text{all headings}}} r_{ijn} p_{ij} f_{ij}(S) \tag{6.18}$$

where F_S is the long term p.d.f of stress range, and f_{ij} is the p.d.f governing the stress range in element (i, j). If we assume F_S has Weibull distribution:

$$p(S) = \frac{h}{q}\left(\frac{S}{q}\right)^{h-1} \exp\left(-\frac{S}{q}\right)^{h} \tag{6.19}$$

Eq. (12) now becomes

$$D = \frac{\nu_0 T}{C} q^m \Gamma(1+\frac{m}{h}) \tag{6.20}$$

where q is the scale factor, h is the shape factor. Accordingly Eq. (18) is simplified as

$$D = \frac{\nu_0 T}{C} \sum_{n=1}^{\text{all load case}} g_n q_n \Gamma(1+\frac{m}{h_n}) \tag{6.21}$$

6.5 Rain-flow cycle counting

Because of the non-linearity in structure-wave system response is usually not normal. In addition it will always have a certain bandwidth. These make the assumption in Eq. (9) vulnerable. To identify stress cycle in such situation, a number of cycle counting methods have been devised A brief list is given by Dowling (1972):

- Peak counting method
- Level crossing method
- Mean crossing peak counting method
- Histogram method
- Range counting method
- Range-mean counting method
- Range-pair counting method

- Rain-flow counting method

Over the years rain-flow and range-pair method have distinguished themselves from the rest. Here rain-flow counting will be looked at in detail.

It is Matsuishi and Endo (1968) who conjured up the rain-flow method. Wirsching (1977) also gave an equivalent definition. In this approach only the closed hysteresis loops in stress-strain path will be identified and treated as stress cycles. Rain-flow method comes in many variations. The algorithm suggested by Downing and Socie (1982) is very popular. It can be performed without prior knowledge of the whole stress/strain history. The same is true of the approach by Clormann and Seeger (1986). As an improvement Hong (1987) proposed a scheme, which can count closed hysteresis loops that are otherwise ignored. The main criticism received by rain-flow technique is the loss of load sequence information. To relieve this, Anthes (1997) proposed a method that is suitable for load sequence model. With reference to the method by Hong, a rain-flow algorithm is summarized as follows:

1) Translate stress history into a series of peaks and troughs starting with zero, store the series in A, and denote its length by L;

2) Set $i = 0, j = 0$;

3) If $(Ai_{+1} > A_i$ .AND. $A_{i+3} \geq A_{i+1}$ .AND. $(A_{i+2} \geq A_i$ .OR. $|A_{i+1}| \geq |A_{i+2}|))$ is true go to (6);

4) If $(A_{i+1} < A_i$ .AND. $A_{i+3} \leq A_{i+1}$ .AND. $(A_{i+2} \leq A_i$ or $|A_{i+1}| \geq |A_{i+2}|)$) is true go to (6);

5) Set $i = i + 1$, if $i > L - 2$, go to (8) else go to (3);

6) Set $j = j + 1$, register stress cycle as S_j = ABS $(A_{i+1} - A_{i+2})$;

7) Delete A_{i+1} and A_{i+2} from series A, set $L = L - 2$, go to (3);

8) If $L \geq 4$ repeat (3) through (7) until no more closed loops can be found, that is when $L < 4$ or L keeps unchanged in the successive two counts.

This algorithm differs from Hong's approach only in the treatment of remnant part (step 8). It is more suitable to be used with random field generation technique in Chapter 2. One example is shown in Fig 6.1.

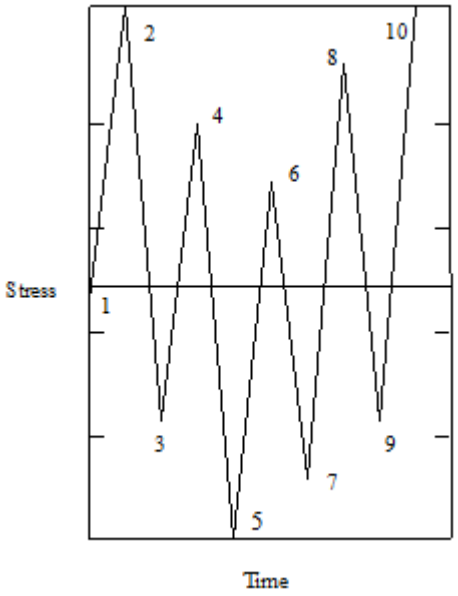


Figure 6.1(a) Cyclic stress history

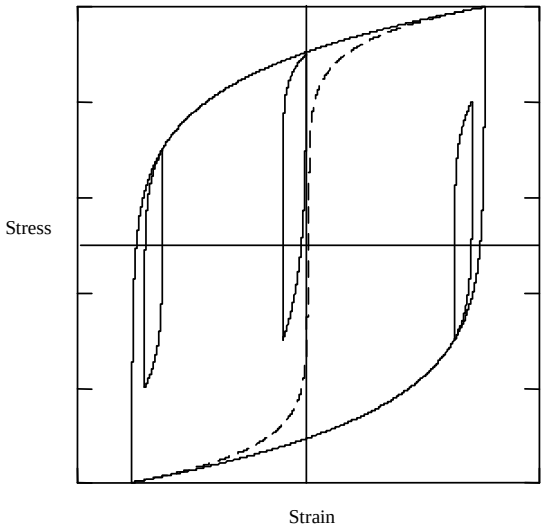


Figure 6.1(b) Closed hysteresis loops

It is worth mentioning that materials are assumed to be cyclic stable in general, which is typical in mild and higher tensile steels in ship hull and marine structures. This assumption reduces substantially the experiment data required to carry out the fatigue analysis.

Strictly speaking none of the parameters in hysteresis loop can be used alone to represent the fatigue damage under cyclic loading. Fatigue criteria in which the number of load cycles prior to fatigue failure is related to either stress or strain range are provisional. More rigorous approach is to represent the fatigue damage in the form of cycle by cycle accumulation of inelastic strain energy until it reaches a critical value. However this is rather intractable in practice.

6.6 Correction of bandwidth and non-Gaussianness

The most pragmatic way to take account of their effects is to offer a correction factor χ to Eq. (10), which yields

$$D = \chi\left(\sqrt{m_0}\right)^m \frac{\nu_0 T}{C} \Gamma(1+\frac{m}{2})(2\sqrt{2})^m = \chi\sigma_x^m D_S \tag{6.22}$$

where σ_x is the standard deviation of stress response, D_S is the fatigue damage of unit σ_x^m. In the meanwhile, using rain flow method we can express the fatigue damage of non-normal and wide band response as

$$D_R = \frac{\nu_R T \mathrm{E}[S_R^m]}{C} = \sigma_x^m \frac{\nu_R T \mathrm{E}[\overline{S}_R^m]}{C} = \sigma_x^m D_{RS} \tag{6.23}$$

where ν_R is the rain flow cycle rate, S_R is the rain-flow stress range, and D_{RS} is the corresponding fatigue damage of unit σ_x^m. S'_R in Eq. (23) is the range of stress standardised by

$$\overline{X} = \frac{X - X_m}{\sigma_x} \tag{6.24}$$

where X_m is the average stress level. Apparent the definition of D_S and D_{RS} is very convenient in random field generation. Now the correction factor can be calculated by

$$\chi = \frac{D_{RS}}{D_S} \tag{6.25}$$

Certainly other counting methods can be used in Eq. (22) as well.

Because of the non-Gaussianness of stress distribution and the complexity of rain-flow counting method, the analytical analysis of Eq. (23) is impossible. So we will apply rain-flow counting to a generated random process and calculate the unbiased estimate of $E[\overline{S}_R^m]$ by

$$\hat{E}[\overline{S}_R^m] = \frac{1}{\hat{N}_R}\sum_{i=1}^{\hat{N}_R}\overline{S}_{Ri}^m \tag{6.26}$$

where $\overline{S}_{Ri}^m$ is the series of rain-flow stress range $\hat{N}_R$ is the corresponding number of rain-flow cycle. According to large number law $\hat{E}[\overline{S}_R^m]$ has an asymptotic normal distribution. The unbiased estimate of its variance is

$$\hat{V}ar[\hat{E}[\overline{S}_R^m]] = \frac{1}{\hat{N}_R - 1}\left[\frac{1}{\hat{N}_R}\sum_{i=1}^{\overline{N}_R}\overline{S}_{Ri}^{2m} - (\hat{E}[\overline{S}_R^m])^2\right] \tag{6.27}$$

In Eq. (27) we assume that $\overline{S}_{Ri}^m$ and $\overline{S}_{Rj}^m$ ($i \neq j$) are statistically independent, since rain-flow counting does not rely on correlation of the original stress series. From Eq. (26) and (27) we can estimate the coefficient of variation of χ by

$$\hat{C}ov[\chi] = \frac{\sqrt{\hat{V}ar[\hat{E}[\overline{S}_R^m]]}}{\hat{E}[\overline{S}_R^m]} \tag{6.28}$$

6.6.1 Effect of bandwidth

To study the effect of bandwidth a Gaussian process is generated. The spectral density is given by Eq. (2.49). To reduce the peak and trough imperfection the time step is chosen as $\Delta t = 1/64$sec. For $m \in [1.10]$, the moment estimates of standardised stress range $\hat{E}[\overline{S}_R^m]$ are shown in Fig. 6.2 against simulation period T. The results of rain-flow counting is shown in Table 6.1, where rain-flow cycle rate is estimated by

$$\hat{\nu}_R = \hat{N}_R / T \tag{6.29}$$

It is found that the convergence rate of moment estimate tends to decrease with the increase of m. To obtain reliable estimate the simulation period must be long enough. In our example, the period T = 262144secs can give well-converged results. This period will be used hereafter. Another thing interesting is that the rain-flow cycle rate ν_R is very close to peak rate n_0. This is also observed by Wirching (1977).

Using the moment estimates by rain-flow method, we can calculate damage D_{RS} by Eq. (23), and then obtain the correction factor by Eq. (25). The result is shown in Fig. (6.3). Apparently, the result from narrow band assumption is rather conservative. The higher the order m, the more conservative it will be.

Another much liked method to take account of the bandwidth is the Dirlik model (Dirlik 1985). Instead of using correction factors it is based on a distribution fitted through the first four moments of the spectral density. This approach will be looked at in spectral fatigue analysis in Chapter 8.

Table 6.1 Results of rain-flow counting

Period T (sec)	Rain-flow cycle rate	Peak rate	Zero-crossing rate
64	0.7969	0.8750	0.5313
128	0.7734	0.8281	0.5469
256	0.8398	0.8594	0.5469
512	0.8672	0.8730	0.5527
1024	0.8633	0.8662	0.5459
2048	0.8594	0.8628	0.5527
4096	0.8628	0.8650	0.5513
8192	0.8542	0.8551	0.5507
16384	0.8588	0.8592	0.5491
32768	0.8617	0.8620	0.5510
65536	0.8588	0.8590	0.5490
131072	0.8585	0.8586	0.5488
262144	0.8587	0.8588	0.5502
Theoretical value: $n_0 = 0.8717,\ \nu_0 = 0.5513$			

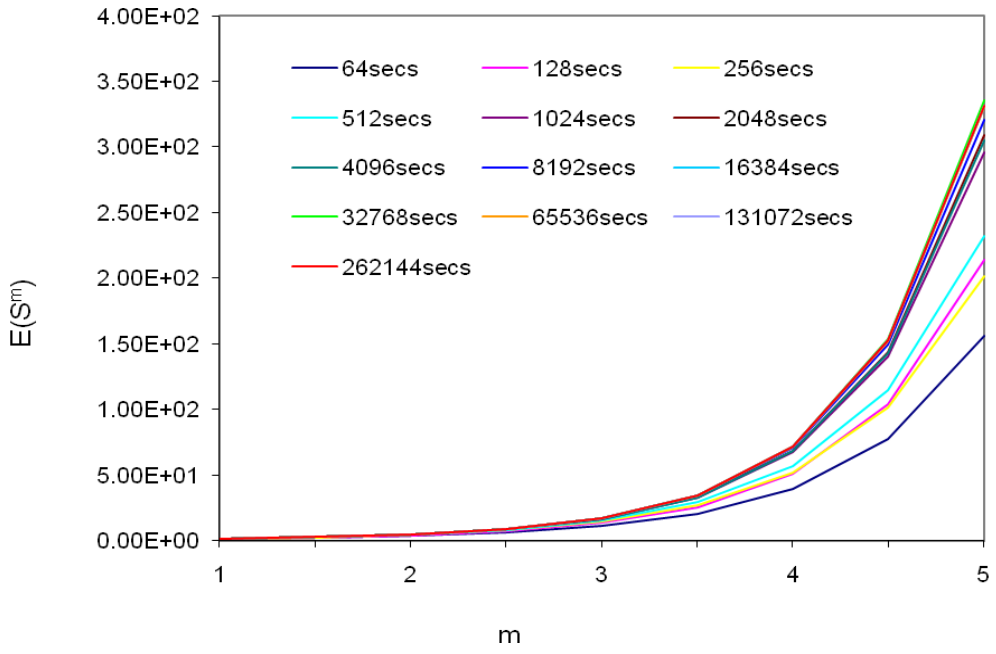


Figure 6.2(a) Moment estimates of stress range by Rainflow counting (m=[1,5])

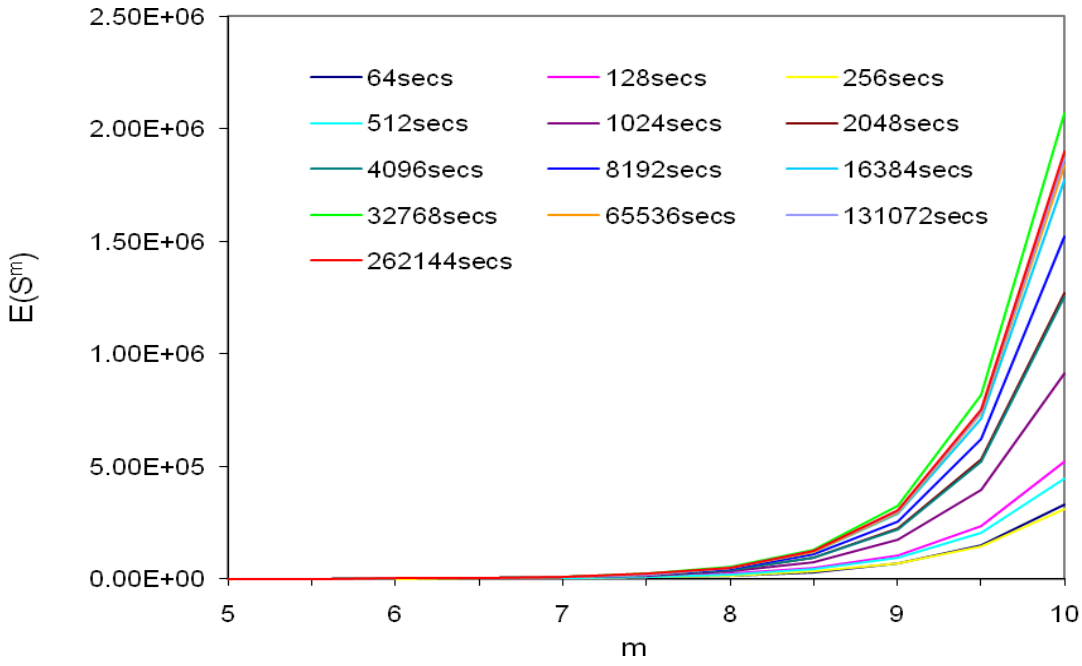


Figure 6.2(b) Moment estimates of stress range by Rainflow counting (m=[5,10])

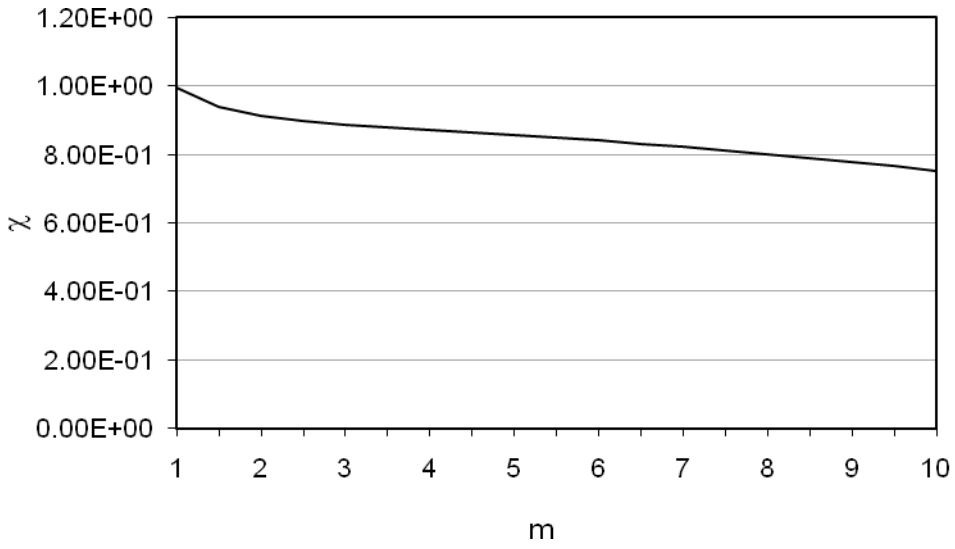


Figure 6.3 Bandwidth Correction Factor

6.6.2 Effect of non-Gaussianness

To generate non-Gaussian random process, the family of distributions by Johnson (1970) is used. It covers a wide range of distributions similar to known ones. The corresponding ZMNL transformations are

$$U = \gamma + \delta \log X \tag{6.30}$$

$$U = \gamma + \delta \log\{X/(1-X)\} \tag{6.31}$$

$$U = \gamma + \delta \sinh^{-1} X \tag{6.32}$$

where γ and δ are shape parameters, U is standard normal variable, and X is stress. Equation (30) corresponds to the family of lognormal, and S_B and S_U are used to denote the distribution defined by Eq. (31) and Eq. (32) respectively. The first four moments of S_U distribution are

$$\mu_1'(X) = -\omega^2 \sinh \Omega \tag{6.33}$$

$$\mu_2(X) = E[(X - \frac{1}{2}(\omega - 1)(\omega \cosh 2\Omega + 1) \tag{6.34}$$

$$\mu_3(X) = -\frac{1}{4}\sqrt{\omega}(\omega - 1)^2\{\omega(\omega + 2)\sinh 3\Omega + 3\sinh \Omega\} \tag{6.35}$$

$$\begin{aligned}\mu_4(X) = \frac{1}{8}(\omega - 1)^2\{&\omega^2(\omega^4 + 2\omega^3 + 3\omega^2 - 3)\cosh 4\Omega \\ &+ 4\omega^2(\omega + 2)\cosh 2\Omega + 3(2\omega + 1)\}\end{aligned} \tag{6.36}$$

where μ'_1 is moment of origin, μ_2, μ_3, and μ_4 are central moments and

$$\begin{aligned}\omega &= \exp(\delta^{-2}) \\ \Omega &= \gamma/\delta\end{aligned} \tag{6.37}$$

Thus we obtain the ZMNL transformation of standardised S_U by

$$\overline{X} = \frac{X - \mu_1'}{\sqrt{u_2}} \tag{6.38}$$

The skewness and kurtosis of S_U family can be calculated by

$$\begin{aligned}\alpha_3 &= \frac{\mu_3}{\mu_2^{3/2}} \\ \alpha_4 &= \frac{\mu_4}{\mu_2\mu_2}\end{aligned} \tag{6.39}$$

They are import parameters to measure the degree of non-normality. For standard normal distribution, we have $\alpha_3 = 0$, $\alpha_4=3$. In Johnson's system a pair of possible values (α_3, α_4) corresponds to only one distribution. Their relationship in S_U family is shown in Fig.6.4 in which $\beta_1=\alpha_3^2$ and $\beta_2=\alpha_4$ are used instead.

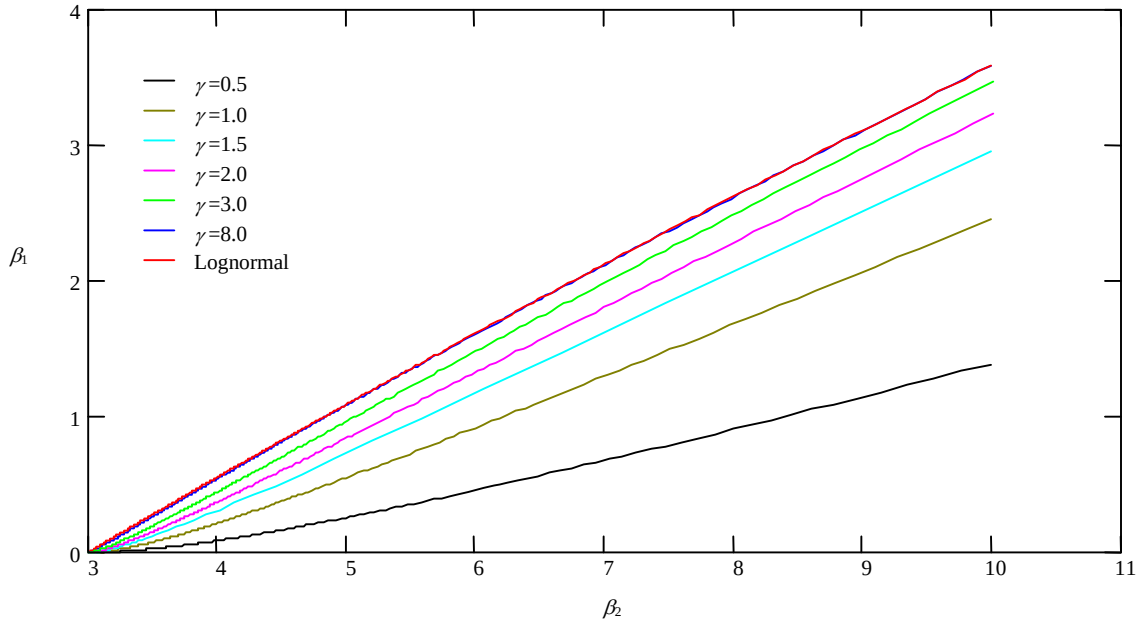


Figure 6.4 Distribution of S_U family

We can learn from Fig. 6.4 that for specific distribution β_1 and β_2 are not independent. In addition, given β_2 there exists a range of possible β_1. The lognormal line in Fig.6.4 is the dividing line of the whole system. Those lines below it belong to S_U, while those above it fall into S_B. Because X in S_B is bounded between (0,1), we will only use S_U and regard lognormal distribution as its asymptotic line (say when $\Omega \geq 8$).

A series of random fields of S_U distribution are generated by Eq. (30). We use parameter γ to control skewness and parameter δ to control kurtosis. The correction factor against β_1, β_2 and m is shown in Fig.6.5.

According to Fig.6.5, both β_1 and β_2 have significant effect on fatigue damage. At $\beta_2 = 3$, the standardised S_U distribution will reduce to normal distribution and the correction factor only reflects the effect of bandwidth. This is the starting point for all the curves in Fig.6.5, since those cases where $\beta_2 < 3$ are exclusively related to S_B family which has little value to be studied. It is very interesting to find that the correction factor is strictly bounded by two curves, one corresponding to symmetric distribution ($\gamma = 0$), the other corresponding to lognormal distribution. They offer the upper and the lower limit of correction factor as β_2 varies from 3 to 10. For the same kurtosis β_2, when γ increase and so is the skewness, the

correction factor will decrease. Furthermore, when m = 1, non-Gaussianness will enhance the effect of bandwidth, and make D_S more conservative. However, when $m \geq 1.5$, non-Gaussianness tends to offset the effect of bandwidth at the very beginning and then outweigh it and make D_S risky. It should be stressed that when we change the sign of γ and hence the sign of skewness, the correction factor will not change. This is due to the simple fact that D_S and D_{RS} are only subject to stress range but not average stress.

The correction approaches in this section stems from the work of Wirsching et al. (1977, 1980), who suggested a coefficient for the effect of bandwidth. Later Lutes et al. (1984) extended the idea to account for the effect of non-normality. In the paper of Winsterstein (1985), an effort is made in analytical approximation. However, none of them employ the FFT technique in simulation. Neither necessary property of generated random field is given in detail. As we have mentioned in Chapter 2, the sample size and parameters of generator have decisive effects on the final results. In addition, no systematic models like Eq. (22) through (31) are used in the previous work to study non-normality. As a result, the effect of skewness is irrationally ignored. Furthermore, there is a common attempt in their work to fit the correction factor as an empirical function either of β_2 and m, or of m and effective bandwidth. But no formula is proposed to include all of those explanatory parameters at one time. This is partly because the complexity involved, partly because the variation of response spectrum used in different areas. In the author's point of view, it is futile to give a panacea, but if we only care about the extremes of correction factor, pragmatic formula may be fitted for specific structural system.

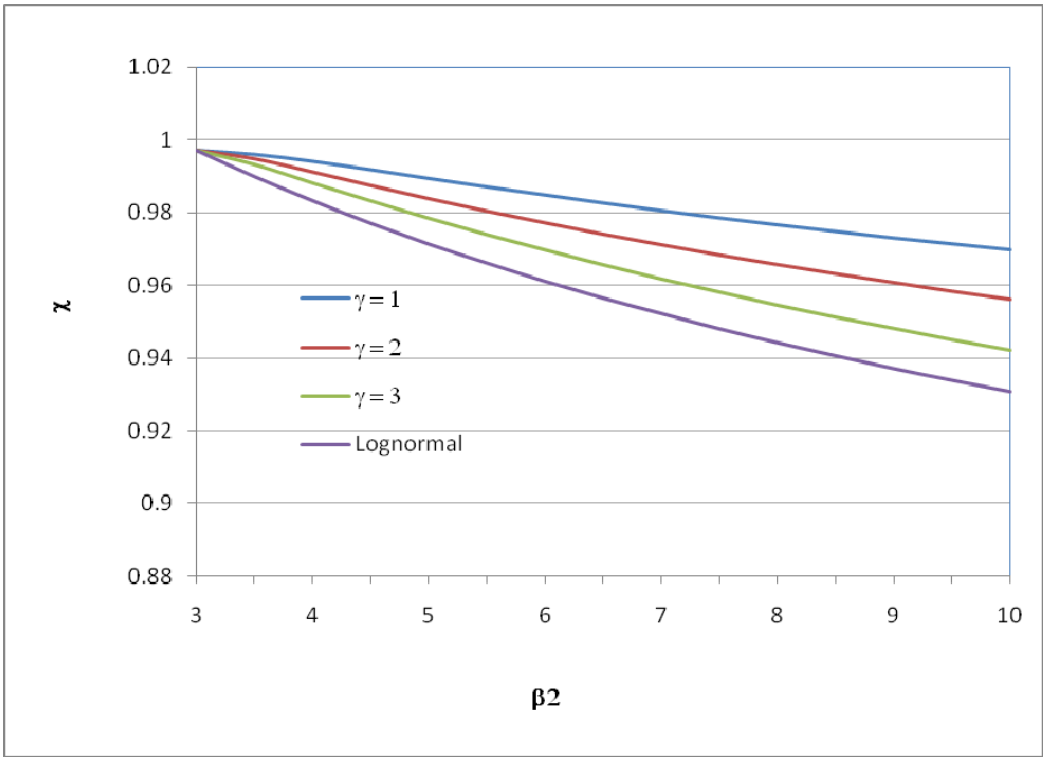


Figure 6.5(a) Correction of non-Gaussianness (m = 1.0)

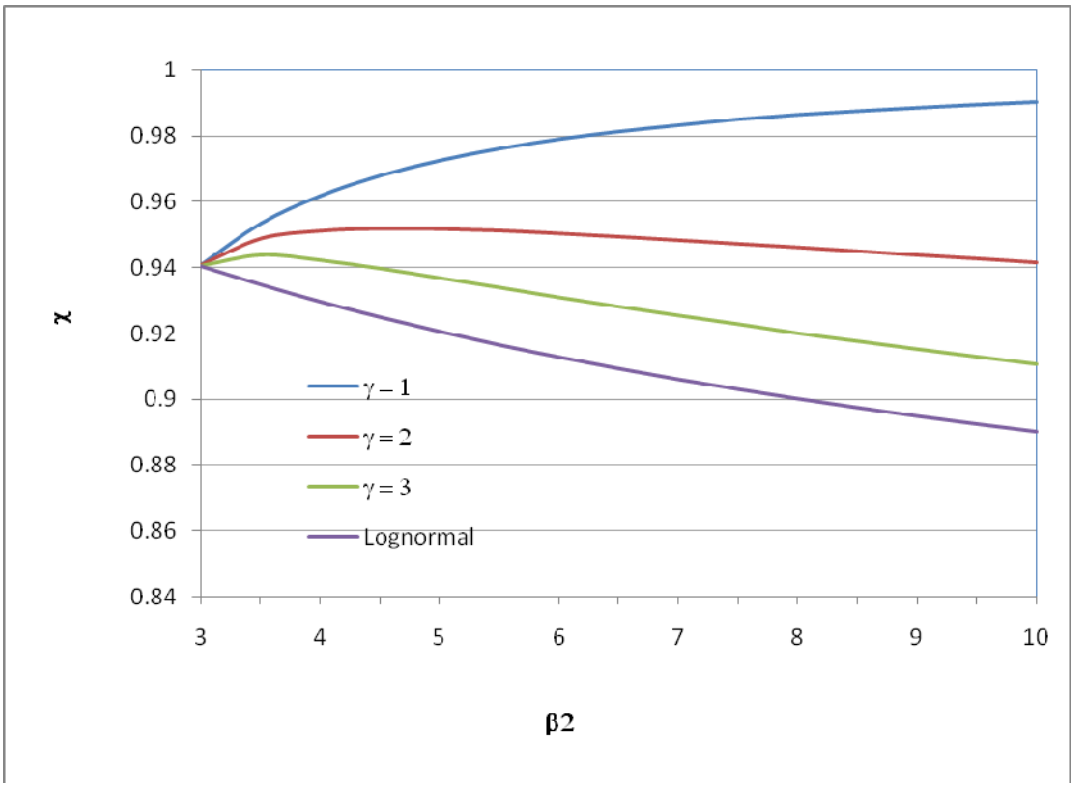


Figure 6.5(b) Correction of non-Gaussianness (m = 1.5)

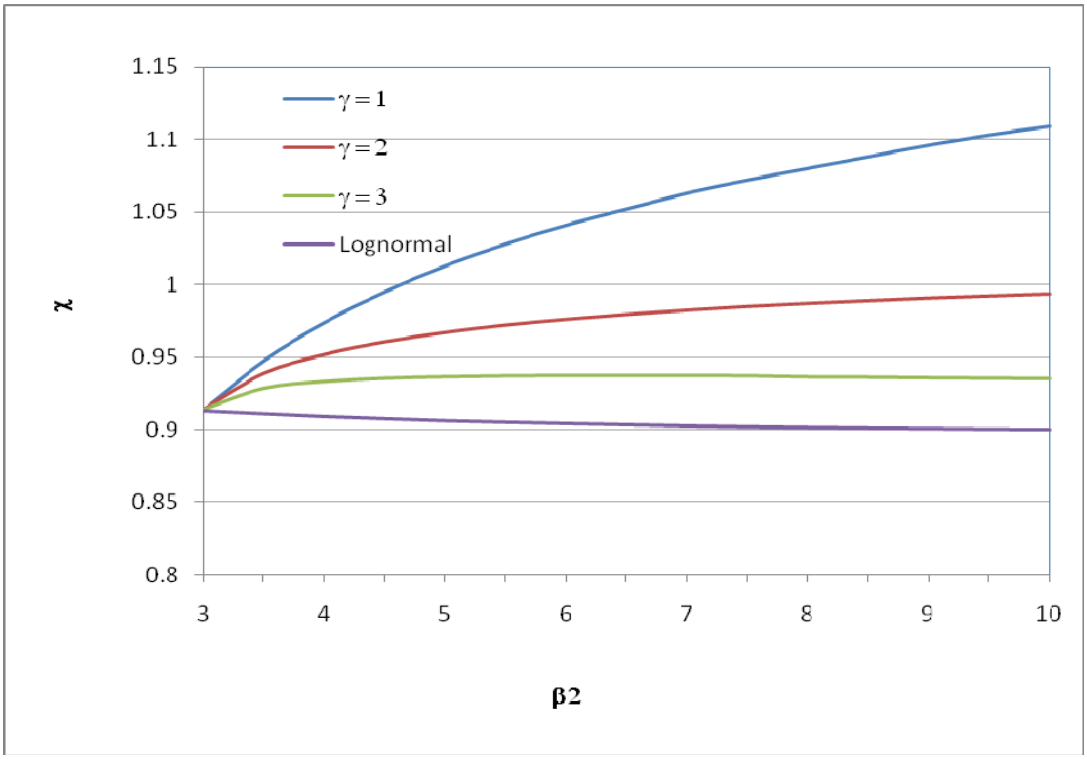


Figure 6.5(c) Correction of non-Gaussianness (m = 2.0)

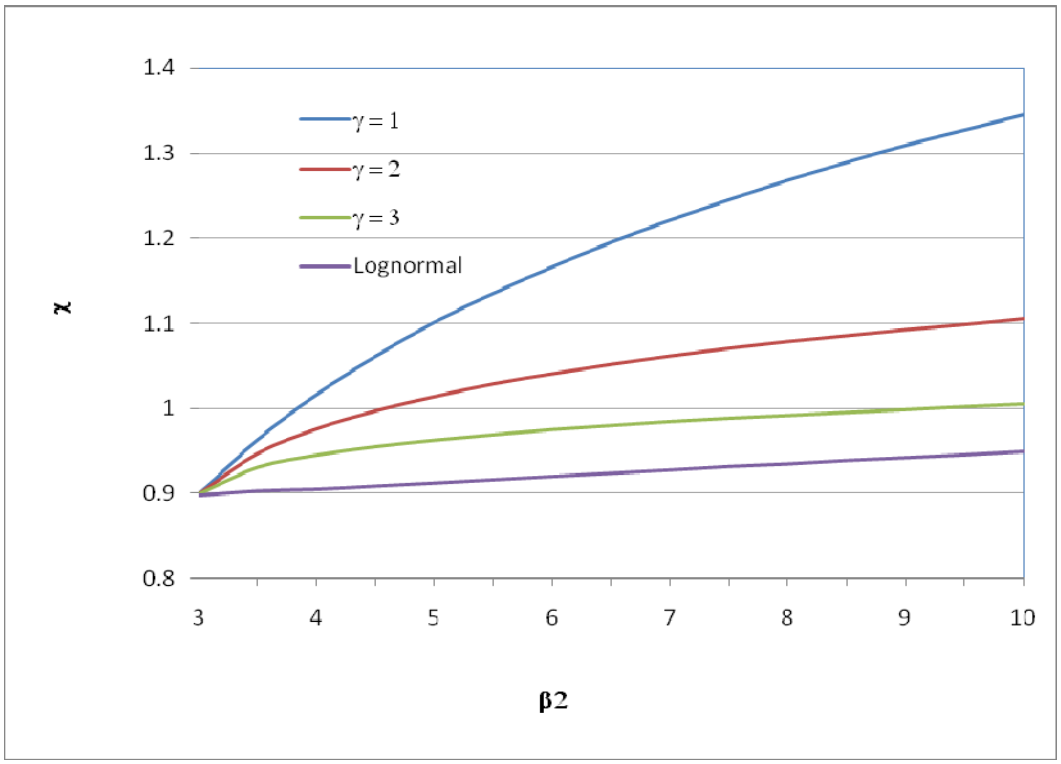


Figure 6.5(d) Correction of non-Gaussianness (m = 2.5)

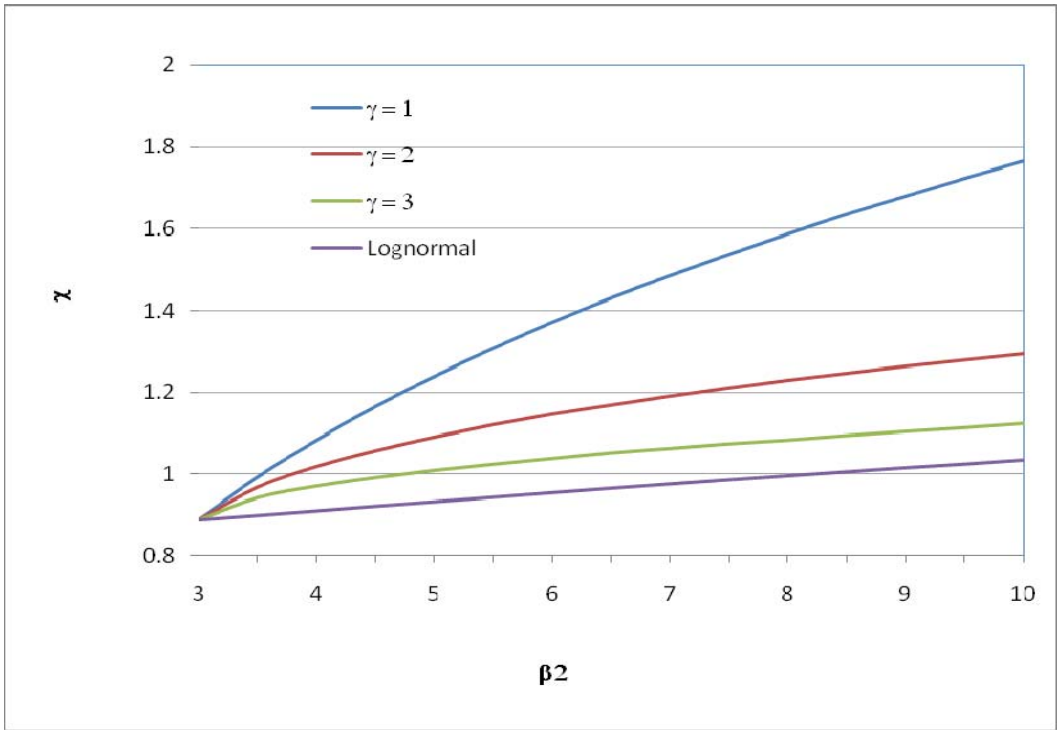


Figure 6.5(e) Correction of non-Gaussianness (m = 3.0)

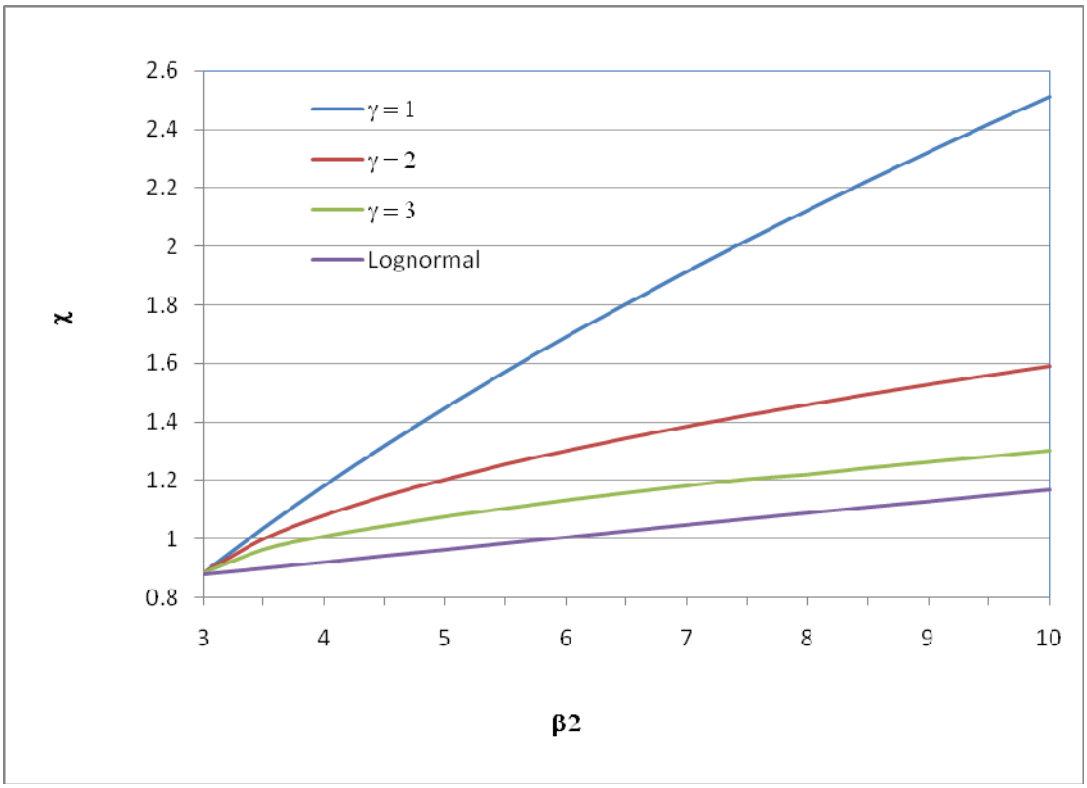


Figure 6.5(f) Correction of non-Gaussianness (m = 3.5)

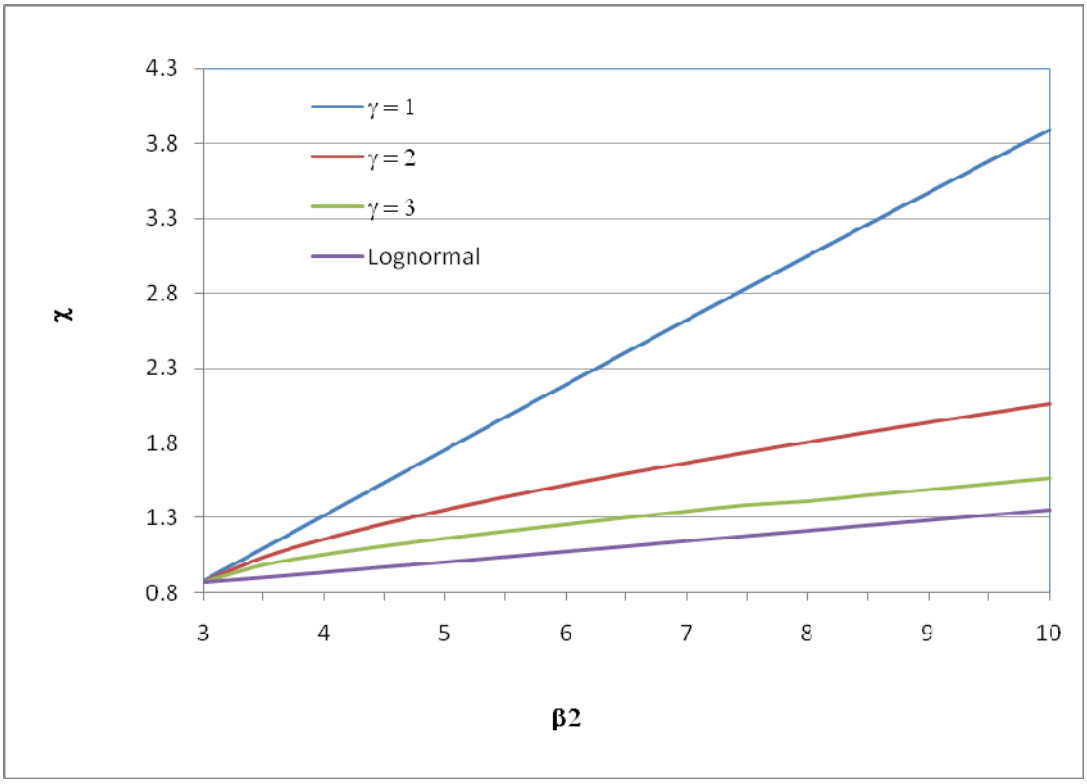


Figure 6.5(g) Correction of non-Gaussianness (m = 4.0)

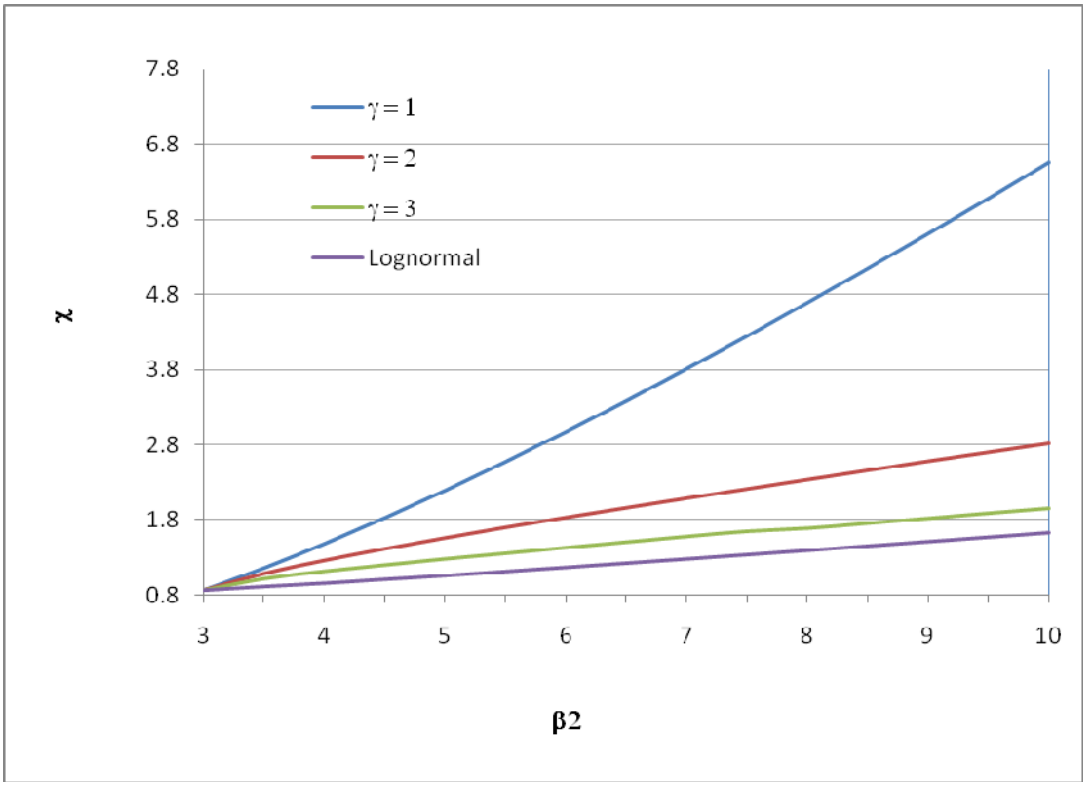


Figure 6.5(h) Correction of non-Gaussianness (m = 4.5)

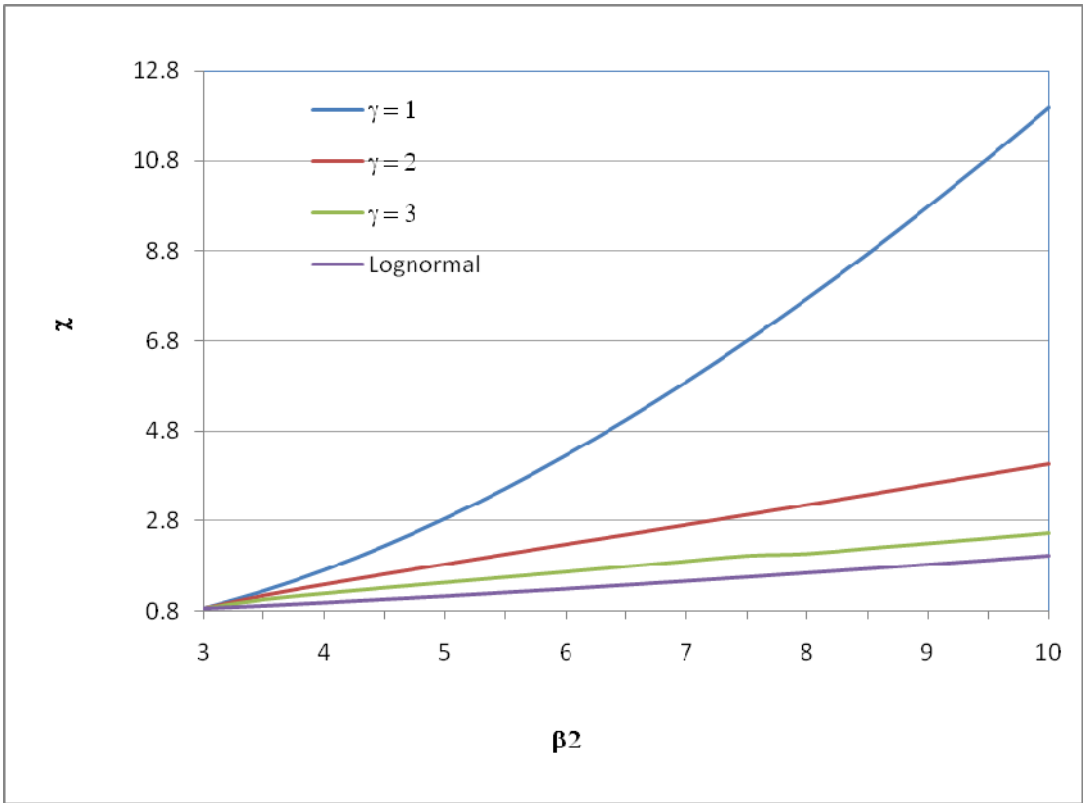


Figure 6.5(i) Correction of non-Gaussianness (m = 5.0)

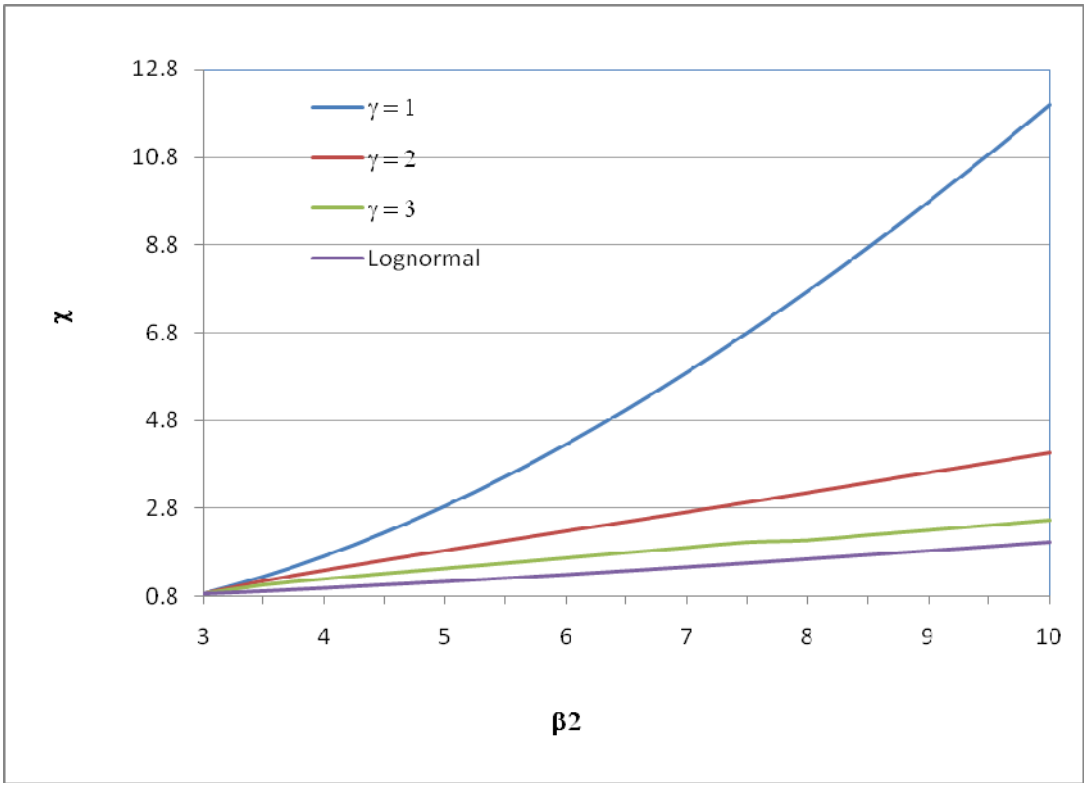


Figure 6.5(j) Correction of non-Gaussianness (m = 5.5)

6.7 Local strain damage model

Coffin's plastic strain criterion based on controlled strain test shows the following relation:

$$\Delta\varepsilon_p = CN^{-\alpha} \tag{6.40}$$

where C and α are material constants. To apply Eq. (40) plastic strain has to be separated from the total strain, which is not a trivial task. Coffin and Tavernelli (1962) extended the relation to cover low and high cycle fatigue:

$$\Delta\varepsilon = \Delta\varepsilon_p + \Delta\varepsilon_e = CN^{-\alpha} + 2\sigma_{-1}/E \tag{6.41}$$

where σ_{-1} is the constant amplitude fatigue limit. Manson's (1962) variation gives

$$\Delta\varepsilon = \Delta\varepsilon_p + \Delta\varepsilon_e = CN^{-\alpha} + BN^{-\beta} \tag{6.42}$$

where C, B, α and β are best fit material constants. Equation (42) allows the cases when $\Delta\varepsilon_e$ exceeds $2\sigma_{-1}/E$ in the hysteresis loop. From Eq. (41) we have:

$$N = C^{1/\alpha}(\Delta\varepsilon - \Delta\varepsilon_e)^{-1/\alpha} = \left(CE/\sigma_{-1}\right)^{1/\alpha}(E\Delta\varepsilon/\sigma_{-1} - 2)^{-1/\alpha} \tag{6.43}$$

Assuming linear damage accumulation as in Eq. (8) the fatigue life can be expressed as:

$$\begin{aligned} N &= \left(\int \frac{p(\Delta\sigma)}{N(\sigma)} d\Delta\sigma\right)^{-1} \\ &= \left(CE/\sigma_{-1}\right)^{1/\alpha} \Big/ \int p(\Delta\sigma)(E\Delta\varepsilon/\sigma_{-1} - 2)^{1/\alpha} d\Delta\sigma \end{aligned} \tag{6.44}$$

Here we assume when fatigue failure occurs $D = 1$.

6.8 Fracture mechanics damage model

Paris (1961) law for fatigue crack propagation states that

$$\frac{da}{dn} = C(\Delta K)^m \tag{6.45}$$

where ΔK is the stress intensity factor range, C and m are material constants. Because ΔK is a characteristic value of the local stress field at the crack tip, Paris model is suitable for non-uniform stress fields or stress concentration zones consisting of multi-component loads and residual stresses. Numerous variations of Eq. (45) have been proposed since. However their applicability is always restricted to certain conditions and none can half match the simplicity in Paris law. Little wonder that it is still predominantly used in practical problems. The fatigue influential factors can well be reflected in the material constant C and m and effective stress intensity factor range.

The generic form of the stress intensity factor range can be written as:

$$\Delta K = SY(a)\sqrt{\pi a} \tag{6.46}$$

where $Y(a)$ is called geometry correction factor. It is dependent on crack geometry, orientation, shape and loading conditions. Substituting Eq. (46) into Eq. (45) gives

$$\int_0^N S^m dn = \int_{a_0}^{a_c} \frac{da}{CY^m(a)(\pi a)^{m/2}} \tag{6.47}$$

Here a_0 is the initial crack length and a_c is the final critical crack length. For constant stress range Eq. (47) becomes:

$$S^m N = \int_{a_0}^{a_c} \frac{da}{CY^m(a)(\pi a)^{m/2}} = Const \tag{6.48}$$

This is the same format as in SN curve approach, which indicates that we could potentially obtain a series of equivalent SN curves based on different assumed crack lengths.

Clearly the central task of LEFM damage model is the calculation of stress intensity factor as the crack propagates. Geometry correction factor *Y(a)* for idealised configurations is available in many handbooks. Background of their derivation can be found in the book by Barsom and Rolfe (1999). For complicated structural configurations experimental method such as photo-elastic tests or numerical method such as FEM is often applied. To represent the crack tip specific element types have been derived with embedded singularity. In linear superimposition technique the original problem is decomposed into a non-singular problem and a singular problem. Only stress fields for the un-cracked body need to be calculated. Path-independent integrals (*J*-integral) derived from energy conservation law can be used to calculate *K* too as in elastic case it is equal to the energy release rate. A review of the relevant numerical methods can be found in the book by Owen and Fawkes (1983).

Empirical formulas for *K* based on numerical calculation and test data are also available. Of which the Newman-Raju (1981) equation is the most referenced by many design codes:

$$K = (\sigma_t + H\sigma_b)F\left(\frac{a}{t}, \frac{a}{c}, \frac{c}{b}, \phi\right)\sqrt{\frac{\pi a}{Q}} \tag{6.49}$$

with

σ_t = remote uniform tension stress

σ_b = remote uniform bending stress

H = function of ϕ, a/c and a/t

a = depth of surface crack

Q = shape factor for elliptical crack

F = stress intensity boundary correction factor

t = plate thickness

c = half length of surface crack

b = half width of cracked plate

ϕ = parametric angle of the ellipse

6.9 Influential factors in fatigue model

The SN curve database is established through testing limited number of baseline specimens subjected to simple cyclic uniform loads. It is unavoidable that the actual structural detail will deviate from the specimens in size, geometry and will operate in different environment and far more complex loading conditions. The effect of influential factors not included in the baseline test can be studied in isolation through specifically designed experiments. In practice engineering judgment is often required to prioritise the factors to be considered in the fatigue model. Note that the effect of these factors should be understood in the context of each fatigue stages.

6.9.1 Stress concentration factor

Stress concentration is generally due to presence of weld (local geometry) and structural detail fabrication (global geometry). The effect of former together with residual stress and properties in heat affected zone is reflected in SN curve to certain extent. Hot-spot stress from FEM analysis is often used to reflect the effect of structural discontinuity at global level (IACS 2006). However in reality the similitude between actual structural detail and test specimen is often questionable. To understand the difference suitable model in the local plastic zone at weld notch root is needed.

For the fatigue crack to initiate not only must the maximum stress range exceed the fatigue limit (if it exists at all) but sufficient inhomogeneous plasticity at the notch root should have developed. The latter means a certain volume of material has to be affected by the cyclic stress above the fatigue limit. Accordingly stress level attained at the notch root must be high enough to drive the microscopic crack in a slip system to break through the first and foremost barrier at a grain boundary (Miller 1993).

The plasticity though at microscopic level will cause redistribution of local stress. As a result the theoretical stress concentration factor based on elastic theory is no longer applicable at notch root. Instead fatigue notch factor is introduced by Peterson (1974) to characterise stress concentration when inelastic material behaviour is involved:

$$K_f = \frac{\sigma_{-1}^m}{\sigma_{-1}^s} \le K_t \tag{6.50}$$

where

σ_{-1}^m = fatigue limit obtained from smooth specimens under fully reversing loads

σ_{-1}^s = fatigue limit of a structural detail under fully reversing loads

K_t = theoretical stress concentration factor

Fatigue notch factor indicates the influence of stress concentration on fatigue life only within the highly localised stress field. This covers stage I and II. Like stress concentration factor, fatigue notch factor is dependent on the material, the loading mode and the geometry of the structural details. The vast variation of these can make its application very complicated. K_f can be used to account for the totality effect of local inelasticity, multi-axial stress, stress gradients and size effect. It can be used in local strain model to consider the mismatch between welded specimen and structural detail (Petinov 1998). Applying notch factor Eq. (44) becomes

$$N = \left(CE/\sigma_{-1}^s\right)^{1/\alpha} \Big/ \int p(\Delta\sigma)(E\Delta\varepsilon/\sigma_{-1}^s - 2K_f/K_t)^{1/\alpha}\, d\Delta\sigma \tag{6.51}$$

6.9.2 Mean stress

The empirical relationship between mean stress and its impact on fatigue limit was first given by Gerber (1874):

$$\sigma_f = \sigma_{-1}\left(1 - \left(\frac{\sigma_m}{\sigma_u}\right)^2\right) \tag{6.52}$$

where

σ_{-1} = fatigue limit obtained under fully reversing loads

σ_m = mean stress

σ_u = ultimate tensile stress

σ_f = corrected fatigue limit

The common understanding that compressive stress can increase the fatigue limit prompted the well known Goodman (1899) relation:

$$\sigma_f = \sigma_{-1}\left(1-\frac{\sigma_m}{\sigma_u}\right) \tag{6.53}$$

These formulas can be used to correct the SN curves with none or partially embedded mean stress effect (in most of the cases a constant stress ratio of $R\in[0,0.1]$ or $R\in[0,0.5]$ is applied in the fatigue test).

Mean stress plays different role in different stage of fatigue damage. In crack initiation stage where the microscopic crack is shear driven the effect of mean stress is not significant because the development of micro-plastic and plastic deformation in this stage is mainly dependent on the stress range. In addition the development of plasticity in stress concentration zone provides conditions for stress relaxation. This is especially true of the cyclical stable materials. Frost (1974) et al. showed that tensile mean stress has no effect on the fatigue strength of steel specimens. Aluminium is more liable to mean stress but to a less extent than may be predicted by those corrections based on Goodman's formula (Goodman 1899). Ebi and Neumann (1990) revealed that short cracks in early stage of crack growth can be open even at fully reversed loading conditions (R = -1) until they reach a certain size. This size which is closely associated with the effect of mean stress can be used as criterion of local fatigue failure. Moreover it was observed that mean stress not only affects the fatigue limit but also the slope of SN curve (Petinov 2003).

In initiation stage mean stress may be studied in a much wider scope. That is it is not necessarily the result of a constant load. In case of variable amplitude random loading the mean stress develops in the transition from one cycle to another, or in every excursion followed by maxima or minima. Rain flow counting may be used to take account of the mean stress of such a nature. The plasticity retardation related to mean stress can be considered by extensive modelling of the stress-strain diagram (Petinov & Yermolaeva 1993).

Ship classification societies have proposed different empirical factors to correct the reference SN curve. Germanisher Lloyd's (2010) correction is as follows:

$$f_R = \begin{cases} 1.0, & \sigma_m \geq \dfrac{\Delta\sigma_{\max}}{2} \\ 1 - c\left(1 - \dfrac{2\sigma_m}{\Delta\sigma_{\max}}\right), & -\dfrac{\Delta\sigma_{\max}}{2} \leq \sigma_m \leq \dfrac{\Delta\sigma_{\max}}{2}, \\ 1 + 2c, & \sigma_m \leq -\dfrac{\Delta\sigma_{\max}}{2} \end{cases} \tag{6.54}$$

where

c = 0, for weld joints subject to constant stress cycles

= 0.15, for weld joints subject to variable stress cycles

= 0.3, for un-welded base material

DNV (2003) proposed reduction factor for the calculated stress range before entering SN curve as shown in Fig. 6.6.

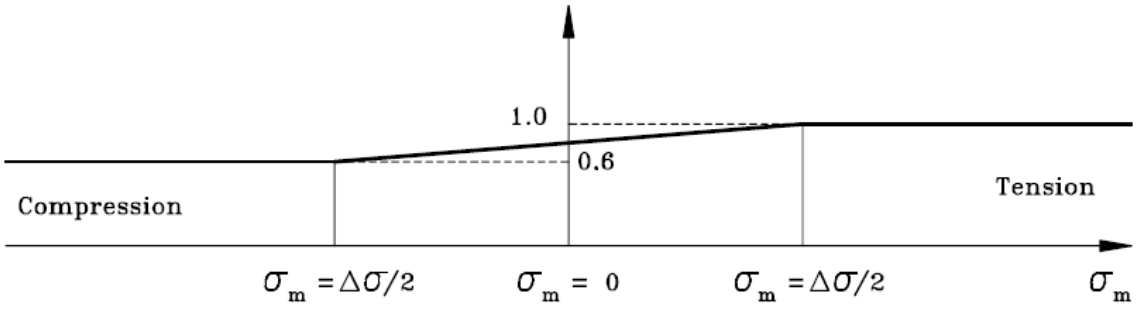


Figure 6.6(a) Reduction factor for base material

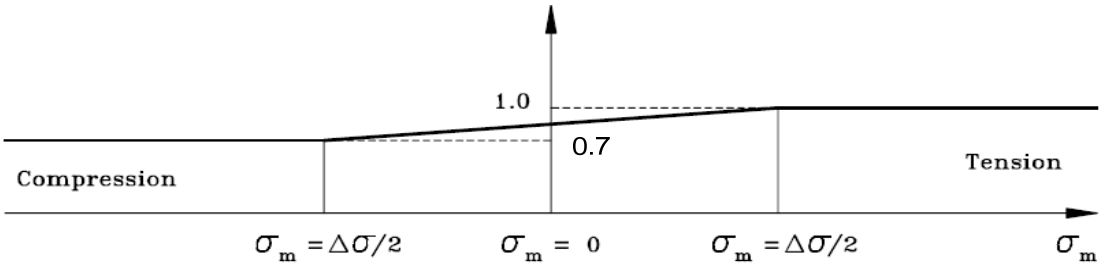


Figure 6.6(b) Reduction factor for welded structures

The CSR rules for double hull tanker recommends reduction factor of 0.6 for compressive pulsating loads i.e. hopper knuckle connection under ballast condition (IACS 2006).

A more sensible model for mean stress effect in the stable propagation stage may be established using LEFM. Elber (1971) proposed the concept of effective stress intensity factor range:

$$\Delta K_{eff} = U\Delta K \tag{6.55}$$

where U is the effective load range ratio taking the flowing form:

$$U = \frac{P_{\max} - P_{op}}{P_{\max} - P_{\min}} \tag{6.56}$$

where P_{max} and P_{min} are the maximum and minimum applied load, P_{op} is the crack opening load. Elber also gave an empirical approximation of U:

$$U = 0.5 + 0.4\frac{K_{\min}}{K_{\max}} = 0.5 + 0.4R, \quad -0.1 \le R \le 0.7 \tag{6.57}$$

6.9.3 Size effect

When testing geometrically similar specimens with different size, it can be observed that smaller specimen tends to produce lower fatigue notch factor. This is called size effect. It indicates that given the same stress range the increase of the sample section and notch toot may end up with decreased fatigue strength. This is because to achieve the same damage level the high-stressed volume should extend relatively deeper in smaller specimen than that in the lager one. A higher stress range in smaller specimen is hence expected (Petinov 2003). The size effect can be exacerbated by applying non-uniform stress in the test, e.g. in cyclic bending or rotation bending. One explanation for size effect is bigger stress concentration zone tends to develop in larger components, which enhances the forming and propagation of macroscopic cracks. In addition the size of plastic zone at the crack tip

also affects the crack opening and closure behaviour during propagation. Size effect is also dependent on the fracture mode.

In ship structures, where model I fracture is common, fatigue strength may decrease with increase of plate thickness due to change of stress gradient, in-homogeneity of micro-damage and through thickness propagation behaviour. Thickness correction factor can be written in the form of

$$f_t = \left(\frac{t_0}{t} \right)^n \tag{6.58}$$

where t_0 is the reference thickness, t is the actual thickness used. Typical values for welded plates from classifications societies and IACS are given in Table 6.2.

Table 6.2 Factors for thickness effect correction

Rules	t_0	n
IACS (CSR BC) (2006)	22mm	1.0 for hatch corner, flat bar or bulb stiffeners, 0.25 otherwise
Lloyd's Register (2004)	22mm	0.1
DNV (2003)	25mm	0.2 for D class SN curve
GL (2010)	25mm	0.17 for welded, 0.1 for toe-ground
ABS (2009)	22mm	0.1 for butt welds ground flush, 0.2 for transverse butwelds, 0.25 for cruciform joints.

6.9.4 Residual stress

Residual stress can considerably affect fatigue crack initiation and propagation. Crack initiation and growth can be accelerated especially after crack nucleation in the weld. Cracks in the compressive field will be arrested. In the tensile field crack may keep growing even under compressive dynamic loads. With the advancement of cracks the

residual stress will redistribute. And the way it redistributes in the test specimen will differ greatly from that in the actual structural details. As a result residual stress will affect the crack path dramatically.

In ship structures two types of residual stresses are present: short range and long range. Short range residual stresses exist either in the weld itself or its proximity. They are caused by thermal contraction of the parts of the cross section of structural member under constraints from the cooler areas. Short range residual stresses are generally of large magnitude. If local plasticity forms under combined constant and dynamic stresses Bauschinger effect will lead to reduced residual stress in cyclical stable or softening materials. This is known as stress relaxation or shake-down. Heat treatment can greatly improve the fatigue strength if the residual stress is tensile. Petinov (2003) showed that FEM method can be used to study the effect of residual stress and believed this is just another highly complicated yet soluble engineering problem.

Long range residual stresses are distributed across the structural member. They are formed during assembly of structural blocks from pre-fabricated parts. Welding shrinkage, local heating and mechanical forces etc. are the common causes. Long range residual stresses are often of moderate magnitude compared to yield stress and exhibit mild gradients. However the constraints at ends may produce significant displacement and strain. Stresses of this nature cannot be relaxed through heat treatment or local plasticity.

6.9.5 Multi-axial stresses

Fatigue failure state under multi-axial stresses is traditionally associated with one of the material strength criteria. For example when the cyclic load is fully reversible and the stress components are fully in-phase we have from the maximum normal principal stress range:

$$\sigma_{-1} = \sigma_x + \sigma_y + \sqrt{\left(\sigma_x - \sigma_y\right)^2 + 4\tau_{xy}^2} \tag{6.59}$$

And the principal shear stress criterion is

$$\tau_{-1} = \sqrt{\left(\sigma_x - \sigma_y\right)^2 + 4\tau_{xy}^2} \tag{6.60}$$

Similar to Von-Mises theory the following combination may be assumed for distortion energy based criterion (Petinov 2003):

$$\sigma_{-1} = \sqrt{\sigma_x^2 + \sigma_y^2 - \sigma_x\sigma_y + 3\tau_{xy}^2 + \sigma_y^2 - \sigma_x\sigma_y + 3\tau_{xy}^2\sigma_x^2} \tag{6.61}$$

The cyclic loads experienced by a ship in service typically consist of external wave pressure, green sea and internal cargo inertia loads. The combination of these components varies in different loading conditions and in different sea states. As a result Eq. (59) - (61) are not applicable here. In practice a simplified criterion based on SN curve and hot-spot stress is often favoured in design. For example in CSR tanker rules stress range normal to the expected crack plane is used for welded hopper knuckle (IACS 2006). By contrast maximum principal stress range is employed in CSR bulk carrier rules (IACS 2006). The principal stresses at the hot spot location having an angle with the assumed fatigue crack greater than 45° are used to calculate the stress range between two extreme load cases. Another approach is to calculate the damage using stresses normal to all possible critical planes between 0° to 360° with pre-defined increment. The maximum damage and the corresponding critical plane calculated in this way will be used in the design (Lloyd's Register 2004).

6.9.6 Corrosion

Corrosion is one of the most common deteriorating factors for ship structures. Corroded plates are more prone to yielding, buckling or fatigue damage. As a matter of fact these damages are often coupled with one being the trigger of another. For example cracks in a plate may change the local boundary condition severely and cause buckling problem. Existing investigations have shown that corrosion affects fatigue strength in more ways than simply reducing the scantlings (Frost et al. 1974, Berge 1976, Burnside 1984,):

- Increased stress level due to scantling deduction
- Pitting corrosion due to material in-homogeneity may form new stress raiser

- Electric polarization may take place with high-stressed area in dissolution acting as the anode and the less stressed surrounding area as cathode
- Protective coating will be damaged leading to more exposure to corrosion
- Decreased barrier for dislocation in favour of the crack initiation
- Hydraulic wedge force may form during crack closure and drive crack propagation
- Fatigue limit demonstrated during test in air may disappear in corrosive environment
- In corrosive environment SN curve shifts considerably to the left when low frequency load is used as opposed to high frequency load.

In order to ensure ship structures have enough strength during its service life corrosion margin must be applied as part of the as designed scantlings. And the plate thickness must be checked during surveys and replaced if the actual diminution is beyond the maximum allowed values (Lloyd's Register 2010).

6.9.7 Material texture

It has long been noticed that material texture is one of the factors contributing to earlier than expected fatigue cracks especially when thickness is used. In the through thickness load transition material texture and sulphur inclusion can reduce the development of plasticity and lead to reduced fatigue resistance. It has been found that specimens cut out perpendicular to the rolling direction demonstrate significant deduction in ultimate stress and fatigue limit compared to those cut out in the rolling direction. Somella (1979) studied lamellar tearing and proposed means to prevent it during construction.

6.10 Reliability formulation of damage model

Fatigue problem is a perfect case to show the two well known sources of uncertainty. On the one hand the sheer number of proposed damage models in the last half century means our knowledge about fatigue is not complete and we have to approach it in a piece-wise manner. Each proposed damage model is a new selective view angle and is subject to epistemic uncertainties. This includes the means of measurement we use in fatigue test. On the other hand for given fatigue model there is inherent natural uncertainties associated

with the identified influential parameters. All this indicates that reliability model is applicable here to improve the safety of fatigue design.

To consider the uncertainty in stress range estimate the actual stress range can be expressed as

$$S_a = BS \tag{6.62}$$

where B is a random modelling error factor. Fatigue damage can now be written as the following generic form:

$$D = \frac{TB^m \Omega}{A} \tag{6.63}$$

where Ω is stress parameter. The Miner's damage criteria is also modified as

$$D \geq \Delta \tag{6.64}$$

Here Δ denotes damage at failure, which is treated as a random variable as opposed to 1 in the deterministic model. In ship structure design required service life is often specified in the codes. So the safety margin can be defined as:

$$Z(\boldsymbol{x}) = T - T_S = \frac{\Delta A}{B^m \Omega} - T_S \tag{6.65}$$

Wirsching (1980) suggested lognormal distribution for all random variables. If we rewrite safety margin as

$$Z = \ln(T) - \ln(T_S) = \ln(\Delta) + \ln(A) - m\ln(B) - \ln(\Omega) - \ln(T_S) \tag{6.66}$$

Apparently Z has Gaussian distribution and the safety index can be obtained by

$$\beta = \frac{\mu_Z}{\sigma_Z} \tag{6.67}$$

where

$$\mu_Z = \mu_{\ln\Delta} + \mu_{\ln A} - m\mu_{\ln B} - \ln\Omega - \ln T_S \tag{6.68}$$

$$\mu_Z = \left(\sigma^2_{\ln\Delta} + \sigma^2_{\ln A} + m^2\sigma^2_{\ln B}\right)^{1/2} \tag{6.69}$$

It can also be seen that T is a random variable with lognormal distribution. Consequently $\ln T$ is a Gaussian random variable with mean value and standard deviation as follows:

$$\mu_{\ln T} = \ln\left(\frac{\tilde{\Delta}\tilde{A}}{B^m\Omega}\right) \tag{6.70}$$

$$\sigma_{\ln T} = \left[\ln\left(1+\delta_T^2\right)\right]^{1/2} = \left\{\ln\left[\left(1+\delta_\Delta^2\right)\left(1+\delta_A^2\right)\left(1+\delta_B^2\right)^{m^2}\right]\right\}^{1/2} \tag{6.71}$$

where δ stand for coefficient of variation. Safety index can now be written as

$$\beta = \frac{\mu_{\ln T} - \ln T_S}{\sigma_{\ln T}} = \frac{\ln\left(\frac{\tilde{T}}{T_D}\right)}{\left[\ln\left(1+\delta^2_{\ln T}\right)\right]^{1/2}} \tag{6.72}$$

Wirsching (1980) suggested that B should be further decomposed as:

$$B = B_M B_S B_F B_N B_H \tag{6.73}$$

where

B_M = Uncertainty due to fabrication and assembly

B_S = Uncertainty due to sea state description

B_F = Uncertainty due to wave load prediction

B_N = Uncertainty due to nominal stress estimation

B_H = Uncertainty due to hot spot concentration factor estimation

Eq. (62) through Eq. (73) are applicable to SN curve, local strain and fracture mechanics damage models. However though the above formulation is easy to follow the modelling uncertainty factor B is not explicitly expressed in terms of design variables. That is to say sensitivity information obtained from such a simplification cannot be directly used to refine the scantlings in the design stage. Other reliability models are needed to complement this.

6.11 Conclusions

- The accuracy of moment estimates of stress range by rainflow counting is decided by both sample length and time step.
- Though Nyquist condition can prevent overlap in frequency domain, it is not enough to prevent the imperfection of peaks and troughs in the generated stress time series. To reduce the imperfection a smaller time step is needed, which will give rise to higher resolution in frequency domain.
- The convergence speed of moment estimates decreases with the increase of order. If the sample size is limited the estimates of higher order is not reliable. On the other hand when a very large sample is used the error of FFT will become significant, and more memory is needed. Out of this consideration, the range of order is chosen from 1 to 5.5 in the chapter. This is enough for most of the engineering problems.
- The rain-flow cycle rate is very close to theoretical peak rate. In practice, the later can be used for the former.
- Though for a practical distribution, skewness and kurtosis are not independent, they all have significant effects on fatigue damage. However, in many previous researches the effect of skewness was ignored improperly.

- Local strain approach is suitable for studying the fatigue behaviour in the plastic zone. Once the crack is beyond original stress concentration zone its applicability will be dramatically weakened. Also to see wider use even in initiation stage existing database must be enhanced for welded structures. This is more suitable for safe-life design concept where tolerance of crack initiation is low.

- Despite the advantages of fracture mechanics it is not suitable for design codes because of the high computational cost. To reasonable estimate NDE (Non Destructive Examination) will have to be used to confirm the initial crack size. This is insurmountable considering the size of the ship. However LEFM is a very good tool in ad hoc damage analysis to determine the residual life of critical locations. Also combined with local strain model it can cover both initiation and propagation stage.

- SN curve approach is still the most practical for design purposes. Years of experience and calibration have made it even more appealing. Just because of its simplicity it can be easily misused by disregarding the similitude between the specimen and real structures. Large scale fatigue test of ship structural details can help stop this loophole.

- Uncertainties in fatigue model call for a reliability model. Fatigue is potentially one of the first areas in traditional ship design where common reliability standards can be established. To further demonstrate the worthiness of reliability approach to the stake holders design variables such as local scantlings need to be part of the reliability model.

References

1. American Bureau of Ships. (2009). *Guidance notes on spectral-based fatigue analysis for vessels*. Huston: ABS.

2. Baker, M. J., (1985). The reliability concept as an aid to decision making in offshore engineering. In J. A. Battjes (Ed.), *Behavior of Offshore Structures* (pp. 75-91). Amsterdam: Elsevier.

3. Barsom, J. M., & Rolfe, S. T. (1999) *Fracture and fatigue control in structures: Application of fracture mechanics* (3rd ed.). Philadelphia: ASTM.

4. Berge, S. (1976). *Corrosion fatigue testing of welding joints at low frequencies* (Report No. SKlR40), Division of Ship Structures, Trondheim: NTH.

5. Burnside, O. H. et al. (1984). *Long-term corrosion fatigue of welded marine steels* (Report No. SSC-326). Ship Structure Committee.

6. Coffin, L. F. (1954) A study of the effects of cyclic thermal stresses in a ductile material. *Transactions of ASME*, 76(6), 931-950.

7. Cramer, E. H., Strathaus, R. S., & Bea, R. G. (1992) *Structural Maintenance for new and existing ships* (Report No. SMP-1-5), Ship Structure Committee.

8. Det Norske Veritas. (2008). *Recommended Practice – Fatigue Design of Offshore Steel Structures* (Report No. DNV-RP-C203). Oslo: DNV.

9. Det Norske Veritas. (2003). *Fatigue assessment of ship structures*. Oslo: DNV.

10. Dirlik, T. (1985). *Application of computers to fatigue analysis*. PHD Thesis, University of Warwick, Warwick.

11. Dowling, N. E. (1972). Fatigue failure predictions for complicated stress-strain histories. *Journal of materials*, JMLSA, **7**(1), 71-87.

12. Ebi, G., & Neumann, P. (1990). Closure behavior of small cracks. *Steel Research*. 61(10), 498-503.

13. Elber, W. (1971) The significance of fatigue crack closure. *Damage and Tolerance in Aircraft Structures*, ASTM STP 485, 230-242.

14. Fatemi, A., & Yang, Y. (1998). Cumulative fatigue damage and life prediction theories: a survey of the state of the art for homogeneous materials. *International Journal of Fatigue*, 20(1), 9-34.

15. Frost, N. E., Marsh, K. J., & Pook L. P. (1974). *Metal Fatigue*. Oxford: Clarendon Press.

16. Gerber, W. (1874). Bestimmung der zulossigen Spannungen in Eisen Constructionen. *Z. Bayer. Arch. Ing. Ver.*, 6, 101-103.

17. Germanisher Lloyd. (2010). *Rules and guidelines*. Hamburg: GL.

18. Goodman, J. (1899). *Mechanics Applied to Engineering*. London: Longmans, Green & Co.

19. Hong, N. (1991). A modified rain-flow counting method. *International Journal of Fatigue*, **13**(6), 465-469.

20. Hsu, T. M. (1988). A simplified method for calculating the remaining fatigue life of cracked structures. OMAE, Houston.

21. Huther, M., Henry J. (1992). *Recommendations for hot-spot stress definition in welded joints* (Report No. XIII-1466-92). IIW.

22. International Association of Classification Societies. (2006). *Common Structural Rules for Double Hull Tankers*. IACS.

23. International Association of Classification Societies. (2006). *Common Structural Rules for Bulk Carriers*. IACS.

24. International Institute of Welding. (1996). *Recommendations on Fatigue of Welded Steel Components* (Report No. XIII-1539-96/XV-845-96). IIW.

25. International Institute of Welding. (1988) *Recommendation on the application of an engineering critical assessment in design, fabrication and inspection to assess the fitness for purpose of welded products*. IIW.

26. Irwin, G. (1957). Analysis of stresses and strains near the end of a crack traversing a plate. *Journal of Applied Mechanics* 24, 361–364.

27. Johnson, N. I., & Kotz, S. (1970). *Distributions in statistics: continuous univariate distributions*. New York: Houghton Mifflin.

28. Lloyd's Register. (2004). *ShipRight Fatigue Design Assessment Level 1 Procedure – Structural Detail Design Guide*. LR.

29. Lloyd's Register. (2010). *Thickness measurement and close-up survey guidance*.

30. Lutes, L. D., et al. (1984). Stochastic fatigue damage Accumulation. *Journal of Structural Engineering*, 110(11), 2585-2601.

31. Manson, S. S. (1953). Behavior of materials under conditions of thermal stress. *Heat Transfer Symposium* (pp. 9-75), University of Michigan Engineering Research Institute.

32. Manson, S. S. (1954). Behavior of materials under conditions of thermal stress. *National Advisory Committee of Aeronautics Technical Note 2933*.

33. Manson, S. S. (1962). Discussion of Ref. 11. *Transactions of ASME, Journal of Basic Engineering*, 84(4), 537-541.

34. Manson, S. S. (1965). Fatigue: a complex subject - some simple approximations. *Experimental Mechanics*, 5(7), 193-226.

35. Matsuishi, M., & Endo, T. (1968). Fatigue of metals subjected to varying stress. Japan Society of Mechanical Engineers, Fukuoka.

36. Madsen, H. O., Krenk, S., & Lind, N. C. (1986). *Methods of Structural Safety*. New Jersey: Prentice-Hall.

37. Miller, K. J. (1993). Materials science perspective of metal fatigue resistance. *Materials Sciencece and Technology*, 9(6), 453-462.

38. Miner, M. A. (1945). Cumulative damage in fatigue. *Journal of Applied Mechanics*, 12, A159-A164.

39. Newman, J. C., & Raju, I. S. (1981). An empirical stress intensity factor equation for the surface crack. *Engineering Fracture Mechanics*, 15(1-2), 185-192.

40. Clormann, U. H., & Seeger, T. (1986). RAINFLOW-HCM, Einzahlverfahren fur Betriebsfestigkeitsnachweise auf werkstoffmechanischer Grundlage, *Stahlbau*, 55, 65-71.

41. Owen, D. R. J., & Fawkes, A. J. (1983). *Engineering fracture mechanics: numerical methods and applications*., Swansea: Pineridge Press.

42. Palmgren, A. (1924). Die Lebensdauer von Kugellagern. *Verfahrenstechinik*, Berlin, 68(14), 339-341.

43. Paris, P. C., Gomez, M. P., & Anderson, W. E. (1961). A rational analytic theory of fatigue. *The Trend in Engineering at the University of Washington*, 13(1), 9-14.

44. Paris, P. C. (1998). Fracture mechanics and fatigue: a historical perspective. *Fatigue and Fracture of Engineering Materials & Structures*. 21(5), 535-540.

45. Peterson, R., E. (1974). *Stress Concentration Factors*: *A Handbook*. New York: J. Wiley & Sons.

46. Petinov, S. V., & Yermolaeva N. S. (1993). Load-history sensitive cyclic curve concept in random load fatigue life predictions. *Ship Technology Research*, 40(2), 117-111.

47. Petinov, S. (1998). The application of S-N curves considering mismatch of stress concentration between test specimen and structure. *Journal of Ship Research*, 42(1), 68-78.

48. Petinov, S. (2003). *Fatigue Analysis of Ship Structures*. New Jersey: Backbone Publishing Company.

49. Somella, J. (1979). *Significance and control of lamellar tearing of steel plate in the shipbuilding Industry* (Report No. SSC-290). Ship Structural Committee.

50. Tanaka, K. (1987). Mechanisms and mechanics of short fatigue crack propagation. JSME, 30(259), l-13.

51. Tavemelli, J. F., & Coffin, L. F. (1962). Experimental support for generalized equation predicting low cycle fatigue. Transactions of ASME, *Journal of Basic Engineering*, 84(4), 533-537.

52. UK Department of Energy. (1990). *Offshore Installations: Guidance on Design, Construction and Certification* (4^{th} ed.). London.

53. Winterstein, S. R. (1985). Non-normal response and fatigue damage. *Journal of Engineering Mechanics*, 111(10), 1291-1295.

54. Wirsching, P. H., & Shehata, A. M. (1977). Fatigue under wide band random stresses using the rain-flow method. *Journal of Engineering Materials and Technology*, 99(3), 205-211.

55. Wirsching, P. H., ASCE, A. M., & Light, M. C. (1980). Fatigue under wide band random stresses. *Journal of the Structural Division*, 106(ST7), 1593-1607.

56. Xu, T. (1997). Fatigue of ship structural details - technical development and problems. *Journal of Ship Research*, 41(4), 318-331.

Chapter 7 Time-variant Reliability Model

7.1 Introduction

The loading on the ship structures is a random process. The strength of the structure system is also subject to time-variant deterioration like corrosion and fatigue damage. In such a context, the probability of failure is equivalent to the first passage probability. That is the probability that structural response surpasses the prescribed safety bounds for the first time within the specified time interval. Despite the extensive study of the first passage problem, closed form solution can only be obtained for very few Markov processes. Blake and Lindsey (1973) have given an overview on the analytical side of level crossing problem. Its variation can be found in the paper by Abrahams (1986). Crandall et al. (1966) studied the problem by simulating the response of a linear one-degree-freedom oscillator excited by Gaussian white noise. The response has proven to be a two dimensional Gaussian-Markov process in the position-velocity space. Their conclusion has been widely quoted as comparison criterion by other researchers. Since decades of intensive investigation have produced massive references in different fields, it is beyond the author to give a complete overview. In this chapter we will only focus on the latest developments that are suitable for numerical treatment. Its application in fatigue reliability is discussed at length.

Most of the existing approaches follow an asymptotic strategy based on Rice's formal representation (Rice 1944) and the assumption of a particular structure in the random process. Gaussianness and stationarity are often used to further simplify the problem. If we assume crossings are independent of each other according to a Poisson process, a simple exponential form can be obtained for the first passage probability. It is asymptotically correct when the second spectral moment gets smaller and the crossing level becomes higher. The error of Poisson assumption hence depends heavily on the bandwidth of the random process. It leads to conservative estimate for narrow band processes and unconservative estimate for wide band processes. Vanmarcke (1975) has developed a

technique allowing for clumping effect to correct Poisson assumption. A single spectra bandwidth parameter is employed in the correction factor. Ditlevsen (1986) associated the Markov structure with the Slepian process instead of the original process and proposed a three parameter distribution model for the duration time of the out-crossing of a wide-banded stationary Gaussian processes. Using the general asymptotic form, Langley (1988) proposed a very efficient formulation for stationary Gaussian process. Given the autocorrelation function, the estimate of mean clump size is greatly simplified.

A big branch of approximations is established by studying point process, which is derived from either the random process itself or its envelope definition. It often involves extensive numerical calculation of the excursion rate in a kind of truncated series. Lin (1970a) examined Rice's in-exclusive series and Kuznetsov's transform from the view of Stratonovich-Kuznetov theory of random points (Stratonovich 1963), which rendered it possible to discretize a continuous process and reduced the computational effort dramatically. It was shown that the first two order truncations of Kuznetsov's series correspond to Poisson and pseudo-Gaussian type of excursion respectively. An explicit formulation was then established upon non-approaching random point theory to take account of higher order cumulants of the excursion rate in Kuznetsov's series. Numerical results showed the non-approaching scheme is more consistent than pseudo-Gaussian and renewal approximation by Rice and Beer (1966). In another study Lin (1970b) proposed a least biased estimate of first passage probability through maximum entropy theory. The constraints used in the maximisation are the first two order moments of the total number of out-crossing in the specified time interval. This approach could be conservative for small times and rather un-conservative for large times. Following the same direction, Yang and Shinozuka (1971) studied the first excursion probability of the narrow-band Gaussian process using discrete extreme point process. In their paper Rayleigh distribution is used for the point process and the joint density function of two consecutive maxima is approximated by the joint density function of two appropriately spaced points on the envelope. Formulations based on Possion and Markov chain assumption are given with a correction of clump size. It is shown that the correlation between up-crossings introduced by Markov process can give good conservative results when the damping coefficients of structural system are sufficiently small which justifies the narrow band assumption. Lin (1972) presented a very suggestive comment on this paper. Complementary to Lin's work (Lin 1970a), Yang and Shinozuka (1972) presented another two approximate solutions

derived from non-approaching random point theory (Stratonovich 1963) and maximum entropy principle. They also carried out the comparison of different approximations, which indicated that the non-approaching random point approximation was the best among the proposed approaches. They pointed out that statistics higher than the second order could play a significant part in the series solution. Soon after that Yang (1972a, 1973) extended the idea to non-stationary case by deriving the joint probability density of a non-stationary envelope, which reduces to Cramer and Leadbetter's definition in stationary case. Markov approach proves distinct from other approximations in such a context. It is emphasised that the effectiveness of each approximation is closely related to the shape correlation function. Krenk (1979) gave a similar discussion on the computational side. In his formulation the initial condition is taken into account and only single integral is involved in numerical calculation. Roberts (1968, 1975) proposed an upper and lower bound of the first passage probability by truncating Rice's series to the third order. Detailed discussion is devoted to the calculation of the out-crossing rates. Thoft-Christensen and Nielsen (1982) showed that the bounds of this kind could be sharpened by using joint crossing rates conditional on previous safe instants in the inclusion-exclusion series.

Shipley and Bernard (1972) and Bernard and Shipley (1972) solved the problem through integral equation governing the first passage density. The formulation was given for a specific stationary Gaussian process, which is a Markov process on a multidimensional position-velocity space. The kernel function is determined by assuming independence between excursions and independence between time and velocity variables. Stressing the integration nature of first passage probability, Madsen and Krenk (1984) generalised the approach by a couple of new kernel approximations. Yang (1975) proposed a recurrence approximation for the first passage probability of Gaussian process. The approach is conservative because of the Markov assumption between excursions. It is equivalent to the discretized version of integral equation method examined by Madsen and Krenk. A alternative integral equation has been used by Roberts (1976) to study the aforementioned linear oscillator system. It resembles the Chapman- Kolmogorov-Smoluchowski equation.

In contrast to integral equation approach, the problem can equally be described by a parabolic partial differentiation equation with certain initial and boundary conditions. Usually Kolmogorov forward or backward equation and Fokker-Planck equation can be obtained based on the theory of diffusion process (Roberts 1986a, 1986b; Bergman and

Heinrich 1982; Lin and Cai 1995). This approach is generally applicable to the one-degree-freedom oscillator systems and requires a lot of computational effort.

7.2 Inclusion-exclusion formulation of first passage probability

First passage probability is defined as the probability of at least one up-crossing of the critical level $b(t)$ by a random process $x(t)$ during time interval [0, t]

$$P_f(t) = 1 - P[n(t) = 0 \mid x(0) < b(0)]P[x(0) < b(0)] \tag{7.1}$$

where $n(t)$ is the number of up-crossing in the interval $(0,t]$. The pdf of the time to the first up-crossing conditional on $x(0) < b(0)$ can be obtained by

$$\begin{aligned} p(t) &= \frac{d}{dt} P_f(t) = p(t \mid x(0) < b(0))P[x(0) < b(0)] \\ &= -\frac{dP[n(t) = 0 \mid x(0) < b(0)]}{dt} P[x(0) < b(0)] \end{aligned} \tag{7.2}$$

An integral equation of $p(t)$ can be established by splitting the crossing rate $\nu(t)$ into two parts (Thoft-Christensen and Nielsen 1982, Madsen 1984):

$$\nu_1(t) = p(t) + \int_0^t K(t \mid \tau)\, p(\tau) d\tau \tag{7.3}$$

where the kernel function $K(t \mid \tau)$ is the crossing rate at t on condition that the first crossing has not happened at τ, namely

$$K(t \mid \tau) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P\{[x(t) < b(t) \mid \tau] \bigcap [x(t + \Delta t) > b(t + \Delta t) \mid \tau]\} \tag{7.4}$$

Likewise we can decompose the joint up-crossing frequency $\nu(t,\ \tau_1),\ \tau_1 < t$, as

$$\nu_2(t, \tau_1) = K(t \mid \tau_1)\, p(\tau_1) + \int_0^{\tau_1} K(t \mid \tau_1, \tau_2)\, p(\tau_1) d\tau_2 \tag{7.5}$$

Substituting Eq. (5) into Eq. (3) yields

$$p(t) = v_1(t) - \int_0^t v_2(t,\tau_1)d\tau_1 + \int_0^t \int_0^{\tau_1} K(t \mid \tau_1, \tau_2)\, p(\tau_2) d\tau_2 d\tau_1 \tag{7.6}$$

Repeating this process gives rise to a formal series representation

$$p(t) = \sum_{j=0}^{\infty} (-1)^j Q_j \tag{7.7}$$

where

$$Q_j(t) = \begin{cases} v_1(t), & \text{for } j = 0 \\ \int_0^t \int_0^{\tau_1} \int_0^{\tau_2} \cdots \int_0^{\tau_{j-1}} v_{j+1}(t, \tau_1, \tau_2, \ldots, \tau_{j-1}, \tau_j) d\tau_1 d\tau_2 \ldots d\tau_{j-1} d\tau_j, & \text{for } j > 0 \end{cases} \tag{7.8}$$

or because of the symmetry of the crossing rate in all its arguments

$$Q_j(t) = \begin{cases} v_1(t), & \text{for } j = 0 \\ \frac{1}{j!} \int_0^t \int_0^t \int_0^t \cdots \int_0^t v_{j+1}(t, \tau_1, \tau_2, \ldots, \tau_{j-1}, \tau_j) d\tau_1 d\tau_2 \ldots d\tau_{j-1} d\tau_j, & \text{for } j > 0 \end{cases} \tag{7.9}$$

The series expressed by Eq. (7) is the well-known inclusion-exclusion series of Rice (1944). For stationary Gaussian random processes it is a non-increasing step function. Upper and lower bounds of $p(t)$ can be obtained upon truncation after an odd or even number since the last term with a kernel integrand is greater than zero. This results in the following inequality

$$R_k(t) \le p(t) \le S_k(t) \tag{7.10}$$

where $k = 0, 1, 2, \ldots$ and

$$R_k(t) = \sum_{j=0}^{2k+1} (-1)^j Q_j \tag{7.11}$$

$$S_k(t) = \sum_{j=0}^{2k} (-1)^j Q_j \tag{7.12}$$

It should be noted that the joint crossing rates and kernel function in Eq. (7) could be conditional on previous safe times, which can lead to sharpened bounds (Thoft-Christensen and Nielsen 1982). If the event of up-crossing is assumed to be independent we can use $v(t)$ for the kernel function in Eq. (3). In this case an explicit solution for $p(t)$ can be obtained

$$p(t) = v_1(t) \exp\left(- \int_0^t v_1(\tau) d\tau \right) \tag{7.13}$$

The corresponding first passage probability is

$$P_f(t) = 1 - P_0 \exp\left(- \int_0^t v_1(\tau) d\tau \right) \tag{7.14}$$

where $P_0 = P[\, x(0) < b(0) \,]$ is the survival probability at starting point. The inclusion-exclusion series for the survival probability in (0,t] is

$$P_s(t) = 1 - P_f(t) = \sum_{j=0}^{\infty} (-1)^j A_j \tag{7.15}$$

where

$$A_j(t) = \begin{cases} 1, & \text{for } j = 0 \\ \dfrac{1}{j!} \int_0^t \int_0^t \int_0^t \cdots \int_0^t v_j(\tau_1, \tau_2, \ldots, \tau_{j-1}, \tau_j) d\tau_1 d\tau_2 \ldots d\tau_{j-1} d\tau_j, & \text{for } j > 0 \end{cases} \tag{7.16}$$

Kuznetsov et.al 1965 have re-expressed Eq. (15) in the form

$$P_s(t) = \exp\left\{ \sum_{j=1}^{\infty} \frac{(-1)}{j!} \int_0^t \int_0^t \int_0^t \cdots \int_0^t g_j(\tau_1, \tau_2, \ldots, \tau_{j-1}, \tau_j) d\tau_1 d\tau_2 \ldots d\tau_{j-1} d\tau_j \right\} \tag{7.17}$$

The relationship between v_j and g_j is analogous to that between moment and cumulant function, thus we have

$$g_1(t_1) = v_1(t_1) \tag{7.18}$$

$$g_2(t_1,t_2) = v_2(t_1,t_2) - v_1(t_1)v_2(t_2) \tag{7.19}$$

$$\begin{aligned} g_3(t_1,t_2,t_3) = {} & v_3(t_1,t_2,t_3) - v_1(t_1)v_2(t_2,t_3) - v_1(t_2)v_2(t_1,t_3) \\ & - v_1(t_3)v_2(t_1,t_2) + 2v_1(t_1)v_2(t_2)v_3(t_3) \end{aligned} \tag{7.20}$$

and etc.

7.3 Approximations of first passage probability

When the time is large the general asymptotic form of the first passage probability is (Kuznetsov et.al 1965)

$$P_f(t) = 1 - A\exp(-\alpha t) \tag{7.21}$$

where A can be interpreted as the probability that $t = 0$ does not fall into a clump and α is known as the limiting decay rate of the first passage probability. Both A and α are functions of the critical level $b(t)$. It can be proved that when $b(t) \to \infty$ we have $A \to 1$ and $\alpha \to v$, which agrees with the results by Cramer and Leadbetter (1967). On the other hand if the autocorrelation function $R(\tau)$ of $x(t)$ is small, implying that $x(t)$ is wide-banded, we also have $\alpha \approx v$. This coincides with the physical argument that successive up-crossings of a wide band process will tend to be independent and will thus constitute a Poisson process. Appropriate truncation of Eq. (17) and assumption of relationship between excursions can lead to different approximations of the first passage probability.

7.4 First passage probability in structural systems

In structural problems the first passage failure probability of random process out-crossings through an uncertain failure surface can be determined by first calculating the first passage

failure probability conditional on an outcome of the time-invariant random variables, and then performing integral over these system parameters to remove the condition. The randomness in a structural system usually consists of time independent random vector $\boldsymbol{X}$ and time-dependent random vector process $\boldsymbol{Z}(t)$. The first passage failure probability in $(0,T)$ can be written as

$$P_f(T) = P\left(\min_{(0,T)} G(\boldsymbol{X}, \boldsymbol{Z}(t), t) \le 0\right) = \int_X f_X(\boldsymbol{x}) P_f(\boldsymbol{x}) d\boldsymbol{x} = E_X\left[P_f(\boldsymbol{X})\right] \tag{7.22}$$

So the centre task here is to determine the conditional failure probability

$$P_f(\boldsymbol{X}) = P\left(G(\boldsymbol{X}, \boldsymbol{Z}(t), t) \le 0\right) \tag{7.23}$$

Similar to Eq. (7) and (15) an upper bound can be obtained as follows

$$P_f(\boldsymbol{X}) \le P_{f_0}(\boldsymbol{X}) + \int_0^T \nu(t, \boldsymbol{X}) dt \tag{7.24}$$

where $P_{f_0}(\boldsymbol{X})$ is the instantaneous failure probability. For narrow-banded process Eq. (24) tends to give quite conservative estimate. An asymptotic approximation similar to Eq. (21) has been given by Ditlevsen (1971) and Vanmarcke (1975):

$$P_f(\boldsymbol{X}) \approx 1 - \left(1 - P_{f_0}(\boldsymbol{X})\right) \exp\left\{-\frac{\int_0^T \nu(t, \boldsymbol{X}) dt}{1 - P_{f_0}(\boldsymbol{X})}\right\} = 1 - \left(1 - P_{f_0}(\boldsymbol{X})\right) \exp\left\{-\frac{\nu(\boldsymbol{X}) T}{1 - P_{f_0}(\boldsymbol{X})}\right\} \tag{7.25}$$

If we further introduce a vector $\boldsymbol{Q}(t)$ to represent long term variations in time such as slowly varying sea state and wind velocity regimes, etc the following bounds will apply (Wen & Chen 1989a, 1989b, 1990; Schall 1990):

$$\begin{aligned} &E_S\left[E_Q\left[1 - \exp\left\{-\int_0^T \nu(t, \boldsymbol{X}, \boldsymbol{Q}) dt\right\}\right]\right] \le P_f(T) \le E_S\left[1 - \exp\left\{-E_Q\left[\int_0^T \nu(t, \boldsymbol{X}, \boldsymbol{Q}) dt\right]\right\}\right] \\ &\le 1 - \exp\left\{-E_S\left[E_Q\left[\int_0^T \nu(t, \boldsymbol{X}, \boldsymbol{Q}) dt\right]\right]\right\} \end{aligned} \tag{7.26}$$

In Eq. (26) the instantaneous probability is ignored for simplicity and independent out-crossings are implied. For stationary crossing problems Næss (1984) gave the following relation:

$$P_f(T) \approx E_S\left[1-\exp\left\{-E_Q\left[\nu(\boldsymbol{X},\boldsymbol{Q})T\right]\right\}\right] \tag{7.27}$$

7.5 Calculation of crossing rate

Through extending Rice's formula (1944) Belyaev (1968, 1969) gave the following conditional mean out-crossing rate:

$$\nu = \nu(\boldsymbol{x},t) = \iint_{\partial G_{\dot{Z}_n} > \partial \dot{G}_n} \left(\dot{z}_n - \partial \dot{G}_n\right) f_{\boldsymbol{Z}\dot{Z}_n}(\boldsymbol{z},\dot{z}_n)\,d\dot{z}_n\,d\boldsymbol{z} = \int_{\partial G} \nu_{local}(\boldsymbol{z})\, f_z(\boldsymbol{z})\,d\boldsymbol{z} \tag{7.28}$$

where

$$\nu_{local}(\boldsymbol{z}) = \int_{\dot{Z}_n > \partial \dot{G}_n} \left(\dot{z}_n - \partial \dot{G}_n\right) f_{\dot{Z}_n|\boldsymbol{Z}}(\dot{z}_n|\boldsymbol{z})\,d\dot{z}_n \tag{7.29}$$

where $\nu_{local}(\boldsymbol{z})$ is the local mean out-crossing rate, $f_{\dot{Z}_n|Z}(\dot{z}_n|\boldsymbol{z})$ is the conditional probability density function of $\dot{Z}_n(t)$ given $\boldsymbol{Z}(t)=\boldsymbol{z}$ at time t, $f_{Z\dot{Z}_n}(\boldsymbol{z},\dot{z}_n)$ is the joint probability of $\boldsymbol{Z}(t)$ and $\dot{Z}_n(t)$.

For scalar process $Z(t)$ and threshold $\xi(t)$ Eq. (28) reduces to Rice's formula:

$$\nu = \int_{\dot{\xi}} \left(\dot{z} - \dot{\xi}\right) f_{\xi\dot{Z}}(\xi,\dot{z})\,d\dot{z} \tag{7.30}$$

For a Gaussian process out-crossing a fixed level ξ we have

$$\nu = \frac{1}{\sqrt{2\pi}} \frac{\sigma_{\dot{Z}}}{\sigma_Z} \varphi\left(\frac{\xi - \mu_Z}{\sigma_Z}\right) \tag{7.31}$$

Closed form for surface integral in Eq. (28) can only be obtained for simple geometry configurations. For other cases numerical integration must be carried out, which is rather prohibitive when dimension of the problem is large. For isotropic time dependent process same linearization as in FORM for time-independent problems can be used. But for non-isotropic process there is no unique optimal linearization point. Directional simulation may be used to determine the out-crossing rate of a vector process from an arbitrary convex safe set (Ditlevsen 1996).

The mean out-crossing rate can be expressed as a sensitivity measure of failure probability of an attendant parallel system (Hagen & Tvedt 1991):

$$\nu = \frac{\partial}{\partial \theta} P\left(\dot{G} < 0 \cap G + \dot{G}\theta < 0\right) \tag{7.32}$$

where

$$\dot{G} = \nabla G(t)\dot{Z}(t) + \frac{\partial G(t)}{\partial t} \tag{7.33}$$

The method is applicable to both Gaussian and non-Gaussian processes, stationary and non-stationary. It should be noted that the parametric sensitivity factor formulation can be applied for any distribution model for twice differentiable continuous variables. In case of multiple limit state function we have:

$$\nu = \sum_{i=1}^{m} \frac{\partial}{\partial \theta_i} P\left(\bigcap_{j \neq i}^{m} \dot{G}_j \leq 0 \cap G_i + \dot{G}_i \theta_i \leq 0 \cap -\dot{G}_i \leq 0 \right)_{\theta_i = 0} \tag{7.34}$$

Relation in Eq. (32) makes it possible to make use of any time-independent method that enables calculation of the parametric sensitivity.

7.6 Nested reliability integration

Once $P_f(\boldsymbol{X})$ is known the unconditional estimate in Eq. (22) can be done through nested reliability integration approach by Wen & Chen (1987).

Introducing auxiliary random variable

$$X_{n+1} = \min_{(0,T)} G(\boldsymbol{X}, \boldsymbol{Z}(t), t) \tag{7.35}$$

we have limit state function

$$X_{n+1} = g(\boldsymbol{X}, X_{n+1}) \tag{7.36}$$

Here $g < 0$ corresponds to failure. Since X_{n+1} depends on $\boldsymbol{X}$ the limit state function is now a function of $\boldsymbol{X}$ implicitly. If we can further transform the problem to another space Y of known distribution through

$$F_Y(Y_{n+1}) = F_{X_{n+1}}(X_{n+1}) \tag{7.37}$$

$$\boldsymbol{Y} = T(\boldsymbol{X}) \tag{7.38}$$

Here T is the transform function, $F_{X_{n+1}}(X_{n+1})$ is the unknown CDF of X_{n+1}. If Y is in the standard U space then T becomes the Rosenblatt transform. Defining limit state function:

$$g'(\boldsymbol{Y}, Y_{n+1}) = Y_{n+1} - F_Y^{-1}\{P_f(T^{-1}(\boldsymbol{Y}))\} \tag{7.39}$$

It can be proved that

$$\begin{aligned} P_f(T) &= \int_{g'(\boldsymbol{Y},Y_{n+1})<0} f_Y(\boldsymbol{Y}) f_{Y_{n+1}}(Y_{n+1}) d\boldsymbol{Y} dY_{n+1} \\ &= \int_{\boldsymbol{Y}} f_{\boldsymbol{Y}}(\boldsymbol{Y}) \left\{ \int_{-\infty}^{F_Y^{-1}\{P_f(T^{-1}(\boldsymbol{Y}))\}} f_{Y_{n+1}}(Y_{n+1}) dY_{n+1} \right\} d\boldsymbol{Y} \\ &= \int_{\boldsymbol{Y}} f_{\boldsymbol{Y}}(\boldsymbol{Y}) P_f(T^{-1}(\boldsymbol{Y})) d\boldsymbol{Y} \\ &= \int_S f(\boldsymbol{X}) P_f(\boldsymbol{X}) d\boldsymbol{X} \end{aligned} \tag{7.40}$$

Reliability methods discussed in the previous chapters can now be applied here to estimate the failure probability of Eq. (39), which can be written as the equivalent form:

$$g'(\boldsymbol{Y}, Y_{n+1}) = F_Y(Y_{n+1}) - P_f\left(T^{-1}(\boldsymbol{Y})\right) \tag{7.41}$$

Kuo and Wirsching (1996) suggested that CDF of time to first passage can be fitted using Eq. (41) and advanced mean value (AMV) scheme.

7.7 Time-variant fatigue reliability

The reliability model given in the end of Chapter 6 already has time T as criteria. It can be viewed as a special case of Eq. (22) and can be solved as time invariant problem. The limit state with the most structural deterioration at T will have to be used. Strictly speaking this gives a lower bound of reliability estimate. But it is believed to be suitable for high cycle fatigue coupled with low speed of deterioration say due to corrosion. A better estimate may be made using Eq. (32).

If fracture mechanics model is used with brittle fracture as the criteria a true time variant model will become necessary as the stress must be treated explicitly as random process in the limit state function:

$$g = R(t, \boldsymbol{X}) - S(t) \tag{7.42}$$

where $\boldsymbol{R}$ is system resistance and S is the stress process. Using the fracture toughness K_c we have

$$R(t, \boldsymbol{X}) = \frac{K_C}{Y(a)\sqrt{\pi a}} \tag{7.43}$$

Where a can be solved implicitly from Eq. (6.48) or explicitly through simplified version of Y(a). Marley and Moan (1992) discussed different approximations of the time variant problem. Environment load was treated as piecewise stationary processes. Both time-variant (TV) and time-invariant (TI) approaches were investigated depending on whether the out-crossing rate in Eq. (25) needs to be calculated. In TI approach X_{n+1} in Eq. (36) was

fitted using Weibull distribution so that limit state in Eq. (41) becomes explicit. Compared to Monte Carlo simulation both gave good enough estimates. But the accuracy in TV version is understandably higher. The Monte Carlo simulation method may be summarised as follows:

Algorithm 7.1

1) Generate a sample for vector $\boldsymbol{X}$ in Eq. (42).

2) Divide sea state in 3 to 12 hour blocks up to the integration limit T. In each block carry out spectral analysis assuming stress response is Gaussian and narrow-banded. So stress range follows Rayleigh distribution.

3) The accumulated damage (system deterioration) can now be calculated in each time block. When evaluating the stress factor $\boldsymbol{\Omega}$ joint distribution of significant wave height and wave period is assumed to be time dependent, such as the one by Haver (1986):

$$f_{H_s}(h)=\begin{cases}\dfrac{\exp\left(-0.5\left[\left(\ln(h)-\lambda\right)/\kappa\right]^2\right)}{\sqrt{2\pi}\kappa h}, & 0\le h\le\eta\\ \beta\rho^{-\beta}h^{\beta-1}\exp\left(-\left(h/\rho\right)^{\beta}\right), & h>\eta\end{cases} \tag{7.44}$$

$$f_{T_p|H_s}(t|h)=\frac{1}{\sqrt{2\pi}\gamma t}\exp\left(-0.5\left[\left(\ln(t)-\mu\right)/\gamma\right]^2\right) \tag{7.45}$$

where parameters in the distribution are:

$$\begin{aligned}&(\eta,\lambda,\kappa,\rho,\beta)=(3.27,0.836,0.6132,2.822,1.547)\\&\gamma^2=0.005+0.085\exp\left(-0.13h^{1.34}\right)\\&\mu=1.59+0.42\ln(h+2)\end{aligned}$$

The use of joint distribution as in Eq. (44) and (45) will prevent the more demanding random process for the slow varying part of the sea state.

4) Calculate out-crossing rate for each time block.

5) Estimate conditional failure probability in Eq. (23).

6) Repeat step 1-5 in proper Monte Carlo simulation. The variance reduction technique discussed in Chapter 3 can be used.

7.8 Conclusions

- Time variant problems can be transferred to nested time-invariant reliability problems. This approach fully leverages development in FORM/SORM, MCS and RSM methods and system reliability theory.
- As already demonstrated transform and inversion plays an important role to convert an intractable model to a soluble one.
- The complexity in evaluating out-crossing rate means Monte Cargo simulation techniques can be a promising choice where system response analysis is not costly.

References

1. Abrahams, J. (1986). A survey of recent progress on level-crossing problems for random processes. In I. F. Blake, & H. V. Poor (Eds.). *Communications and Networks*: *A Survey of Recent Advances* (pp. 7-25), New York: Springer-Verlag.

2. Belyaev, Y. K. (1968). On the number of exits across the boundary of a region by a vector stochastic process. *Theory of Probability Applications*, 13, 320-324.

3. Belyaev, Y. K., & Nosko, V. P. (1969). Characteristics of excursions above a high level for a Gaussian process and its envelope. *Theory of Probability Applications*, 14, 296-309.

4. Bernard, M. C., & Shipley, J., W. (1972). The first passage problem for stationary random structural vibration. *Journal of Sound and Vibration*, 24(1), 121-132.

5. Blake, I. F., & Lindsey, W. C. (1973). Level crossing problems for random processes. *IEEE Transactions on Information Theory*, IT-19, 295-315.

6. Crandall, S. H., Chandiramani, K. L., & Cook, R. G. (1966). Some first passage problems in random vibration. *Journal of Applied Mechanics*, 33(3), 532-538.

7. Cramer, H., & Leadbetter, M., R. (1967). *Stationary and related stochastic processes.* New York: John Wiley & Sons.

8. Ditlevsen, O. (1971). *Extremes and first passage times with applications in civil engineering: some approximate results in the theory of stochastic process*. PHD thesis, Lyngby, Denmark: Technical University of Denmark.

9. Ditlevsen, O, (1986). Duration of visit to critical set by Gaussian process. *Probabilistic Engineering Mechanics*, 1(2), 82-93.

10. Ditlevsen, O. & Madsen H. O. (1996). *Structural reliability methods*, Chichester: John Wiley and Sons.

11. Hagen, Ø., & Tvedt L. (1991). Vector process out-crossing as parallel system sensitivity measure. *Journal of Engineering Mechanics*, 117(10), 2201-2220.

12. Krenk, S. (1979). Non-stationary narrow-band response and first-passage probability. *Journal of Applied Mechanics*, 46(4), 919-924.

13. Haver, S. & Nyhus, K.A. (1986). A wave climate description for long term response calculations. *Proceedings of Ocean Mechanics and Arctic Engineering*.

14. Kuo C.-J., & Wirsching P. H. (1996). Fatigue reliability analysis based on time dependent first passage. In D. M. Frangopol, M. D. Grigoriu (Eds.). *Probabilistic Mechanics and Structural Reliability*. *Fatigue Reliability Analysis* (pp. 466-469). New York: ASCE.

15. Kuznetsov, P., I., Stratonovitch, R., L., & Tikhonov, V., I. (1965). On the duration of excursions of random functions. In P. I. Kuznetsov, R. I. Stratonovitch, & V. I. Tikhonov (Eds.). *Non-linear Transformation of Stochastic Processes* (pp. 341-353), Oxford : Pergamon.

16. Langley, R., S. (1988). A first passage approximation for normal stationary random processes. *Journal of Sound and Vibration*, 122(2), 261-275.

17. Lin, Y., K. (1970a) On first excursion failure of randomly excited structures. *AIAA Journal*, 8(4), 720-725.

18. Lin, Y., K. (1970b). On first excursion failure of randomly excited structures, II. *AIAA Journal*, 8(10), 1888-1890.

19. Lin, Y., K. (1972). Discussion: 'On the first excursion probability in stationary narrow-banded vibration'. *Journal of Applied Mechanics,* ASME, 39(2), 632.

20. Lin, Y. K. & Cai, G. Q. (1995). *Probabilistic Structural Dynamics-Advanced Theory and Applications*, Singapore: McGraw-Hill.

21. Lyon, R., H. (1961). On the vibration statistics of a randomly excited hard-spring oscillator II. *Journal of the Acoustical Society of America*, 33(10), 1395-1403.

22. Madsen, P., H., & Krenk, S. (1984). An integral method for the first passage problem in random vibration. *Journal of Applied Mechanics*, ASCE, 51(3), 674-679.

23. Madsen, H., O., Krenk, S., & Lind, N., C. (1986). *Methods of structural safety*, New Jersey: Prentice-Hall.

24. Marley, M. J., & Moan T. (1992). Time variant formulation for fatigue reliability. *Proceedings of Ocean Mechanics and Arctic Engineering*, 2, 161-168.

25. Næss, A. (1984). Technical note: On the long-term statistics of extremes. *Applied Ocean Research*, 6(4), 227-228.

26. Rice S., O. (1944). Mathematical analysis of random noises. *Bell System Journal*, 23, 282-332. 24, 46-156.

27. Rice, J. R., & Beer, F. P. (1966). First occurrence time of high level crossings in a continuous random processes. *Journal of the Acoustical Society of America*, 39(2), 323-335.

28. Roberts, J. B. (1968). An approach to the first passage problem in random vibration. *Journal of Sound and Vibration*, 8(2), 301-328.

29. Roberts, J. B. (1975). Probability of first passage failure for non-stationary random vibration. *Journal of Applied Mechanics*, 42(3), 716-720.

30. Roberts, J. B. (1976). First passage time for the envelope of a randomly excited linear oscillator. *Journal of Sound and Vibration*, 46(1), 1-14.

31. Roberts, J. B. (1986a). First passage probabilities for randomly excited systems: diffusion methods. *Probabilistic Engineering Mechanics*, 1(1), 66-81.

32. Roberts, J. B. (1986b). First passage time for randomly excited non-linear oscillators. *Journal of sound and vibration*, 109(1), 33-50.

33. Schall, G., Faber, M. H., & Rackwitz, R. (1990). Investigation of the ergodicity assumption for sea states in the reliability assessment of offshore structures. *Proceedings of 9th International Conference on Offshore Mechanics and Arctic Engineering*: Vol. II (pp. 1-6). Houston, Texas: OMAE.

34. Shipley, J., W. & Bernard, M. C. (1972). The first passage time problem for simple structural systems. *Journal of Applied Mechanics*, 39(4), 911-917.

35. Stratonovich, R., I. (1963). *Topics in the Theory of Random Noises* (Vol. 1). New York: Gordon and Breach.

36. Thoft-Christensen, P., & Nielsen, S. R. K. (1982). Bounds on the probability of failure in random vibration. *Journal of Structural Mechanics*, 10(1), 67-91.

37. Vanmarcke, E. H. (1975). On the distribution of the first-passage time for normal stationary random processes. *Journal of Applied Mechanics*, ASME, 42(2), 215-220.

38. Wasan, M., T. (1994). Stochastic Processes and Their First Passage Times, *Queen's papers in pure and applied mathematics* (No. 96), Kingston, Ontario: Queen's University.

39. Wen, Y. K., & Chen H. C. (1987). On fast integration for time variant reliability. *Probabilistic Engineering Mechanics*, 2(3), 156-162.

40. Wen, Y. K., & Chen H. C. (1889a) System reliability under time varying loads: I. *Journal of Engineering Mechanics*, ASCE, 115(4), 808-823.

41. Wen, Y. K., and Chen H. C. (1889b) System reliability under time varying loads: II. *Journal of Engineering Mechanics*, ASCE, 115(4), 823-839.

42. Wen, Y. K. (1990). *Structural load modeling and combination for performance and safety evaluation*, Amsterdam: Elsevier.

43. Yang, J.-N. (1972a). Nonstationary envelope process and first excursion probability. *Journal of Structural Mechanics*, 1(2), 231-248.

44. Yang, J.-N. (1972b) Simulation of random envelope processes. *Journal of Sound and Vibration*, 21(1), 73-85.

45. Yang, J.-N. (1973). First excursion probability in nonstationary random vibration. *Journal of Sound and Vibration*, 27(2), 165-182.

46. Yang, J.-N. (1975). Approximation to first passage probability. *Journal of the Engineering Mechanics Division*, ASCE, 101(EM4), 361-372.

47. Yang, J.-N., & Heer, E. (1971). Reliability of randomly excited structures. *AIAA Journal*, 9(7), 1262-1268.

48. Yang, J.-N., & Shinozuka, M. (1971). On the first excursion probability in stationary narrow band vibration. *Journal of Applied Mechanics*, ASME, 38(4), 1017-1022.

49. Yang, J.-N., & Shinozuka M. (1972). On the first excursion probability in stationary narrow band vibration, II. *Journal of Applied Mechanics*, ASME, 39(3), 733-738.

Chapter 8 Spectral Fatigue Reliability of Ship Structures

8.1 Introduction

Traditional spectral fatigue analysis is applicable only to linear systems. Besides, it is often assumed that stress response is narrow-banded and Gaussian. Thus Rayleigh distribution can be used to describe the short term stress range. All this can be rendered vulnerable by the ship-wave system in the real world. For one thing, stress response as well as wave excitation always has certain band width. For another, the stress response is not linearly related to wave height due to non-harmonic wave pressure in the intermittent wetting area and motion-induced inertia loads. As a result the stress response is no longer Gaussian.

To close the first loophole Wirsching & Shehata (1977) proposed a correction factor obtained through rain-flow counting. Its ease of use soon won the method a great deal of popularity in ship and offshore industry. Recently the model originated by Dirlik (1985) has proved more appealing not least because of its consistent performance in a wide range of circumstances (Sherratt et al. 2005). In this model probability density function (PDF) of rain-flow stress range is constructed as a function of moments of spectral density. However it does not remove the Gaussian assumption regarding stress amplitude. To address the non-linear effect Winterstein (1988) came up with a narrow-banded damage model in terms of variance and kurtosis of the stress response. Later Hansen (1994) pioneered an approach based on Longuet-Higgins distribution. It shows an equivalent transfer function of non-linear wave responses can be obtained by time step analysis, in which random sinusoidal wave is applied to the ship structure quasi-statically. In other words the wave excitation in this model is still narrow-banded. Folsø (1998) elegantly adapted this approach for spectral fatigue damage of side shell structures, employing pressure RAOs from linear strip theory.

Despite the above efforts a full-blown fatigue model in frequency domain that can take account of both band-width and non-linearity effects has yet to emerge. The common practice seems to favour the opinion that appropriate linearization plus good treatment of bandwidth effect can give reasonable prediction in fatigue design. A pseudo-excitation approach is proposed in this chapter along this line and applied in reliability case study of real ship structures with damage data. The approach can leverage existing frameworks of spectral fatigue analysis in the industry.

8.2 Short and long term damage

As shown in Chpter 6 the short term fatigue damage in given sea state can be simulated by Miner-Palmgren linear damage accumulation model

$$D = T\nu \int \frac{p(S)}{N(S)} dS \tag{8.1}$$

where D is the fatigue damage index, $\Delta\sigma$ is stress range, $p(S)$ is PDF of the stress range, $N(S)$ is the number of stress cycles that cause fatigue failure under S, T is total service time, and ν is zero crossing rate. Put in generic form Eq. (1) becomes

$$D = \frac{T\nu}{K} E[(S)^m] = \frac{T\Omega}{K} \tag{8.2}$$

where m is inverse slope of SN curve, and K is the intercept. In ship fatigue design two slope SN curve is usually adopted:

$$N(S) = \begin{cases} K_1 S^{-m}, S \geq S_0 \\ K_2 S^{-m'}, S < S_0 \end{cases} \tag{8.3}$$

where m, K_1, m' and K_2 are the inverse slopes and intercepts below and above 10^7 stress cycles respectively; S_0 is the stress range corresponding to the 10^7 cycles in the SN curve.

Under assumption that the stress spectrum is Gaussian and narrow-banded the stress range has Rayleigh distribution and we arrive at:

$$D = \frac{T\nu}{K_1}\left\{\left(2\sqrt{2}\right)^m \sigma^m D_1 + \left(2\sqrt{2}\right)^{m'} \sigma^{m'} D_2\right\} \tag{8.4}$$

with

$$\begin{aligned} D_1 &= \Gamma\left(1+\frac{m}{2}\right) - \gamma\left(1+\frac{m}{2}, \frac{S_0^2}{8\sigma^2}\right) \\ D_2 &= S_0^{m-m'}\Gamma\left(1+\frac{m'}{2}, \frac{S_0^2}{8\sigma^2}\right) \end{aligned} \tag{8.5}$$

where Γ and γ are Gamma and incomplete Gamma function respectively; σ is standard deviation of the stress process. The statistics of stress response can be obtained through the moments of its spectrum $S_y(\omega)$ defined as

$$m_n = \int_0^\infty \omega^n S_y(\omega) d\omega \tag{8.6}$$

The standard deviation can now be calculated by

$$\sigma = \sqrt{m_0} \tag{8.7}$$

The zero-crossing rate (a very good estimate of the peak rate under narrow-band assumption) can be expressed as:

$$\nu = \frac{1}{2\pi}\sqrt{\frac{m_2}{m_0}} \tag{8.8}$$

The bandwidth of the stress process is defined as follows:

$$\varepsilon = \sqrt{1 - \frac{m_2^2}{m_0 m_4}} \tag{8.9}$$

Wirsching (1977) proposed the following correction factor to consider band-width effect:

$$\lambda(m, \varepsilon) = c_a + (1 - c_a)(1 - \varepsilon)^{c_b} \tag{8.10}$$

where

$$\begin{aligned} c_a &= 0.926 - 0.033m \\ c_b &= 1.587m - 2.323 \end{aligned} \tag{8.11}$$

Dirlik's (1985) distribution for stress range has demonstrated consistent performance in wide circumstances:

$$p(S) = \frac{\frac{D_1}{Q} e^{\frac{-Z}{Q}} + \frac{D_2 Z}{R^2} e^{\frac{-Z^2}{2R^2}} + D_3 Z e^{\frac{-Z^2}{2}}}{2\sqrt{2m_0}} \tag{8.12}$$

with

$$\begin{aligned} &Z = \frac{S}{2\sqrt{m_0}}, \alpha = \frac{m_0}{\sqrt{m_2 m_4}}, x_m = \frac{m_1}{m_0}\sqrt{\frac{m_2}{m_4}} \\ &D_1 = \frac{2(x_m - \alpha^2)}{1 + \alpha^2}, R = \frac{\alpha - x_m - D_1^2}{1 - \alpha - D_1 + D_1^2}, D_2 = \frac{1 - \alpha - D_1 + D_1^2}{1 - R} \\ &D_3 = 1 - D_1 - D_2, Q = \frac{1.25(\alpha - D_3 - D_2 R)}{D_1}, \end{aligned} \tag{8.13}$$

The short time damage will now have to be evaluated through numerical integration.

The long term fatigue damage can be written as

$$D = \sum_{L}\sum_{V}\sum_{\theta}\sum_{H_S}\sum_{T_Z} D(L,V,\theta,H_s,T_z)P(L,V,\theta,H_s,T_z) \tag{8.14}$$

where L is loading condition, V is ship speed, θ is ship to wave angle, H_s is significant wave height, T_z is average wave period of zero-upcrossings, and $P(L,V,\theta,H_s,T_z)$ is joint probability of given service condition and sea state, which is usually determined by voyage simulation.

8.3 Pseudo-excitation spectral analysis

8.3.1Spectral analysis for linear system

The stress spectrum of a linear ship structural system subject to stationary wave excitation can be written as

$$S_y(\omega) = S_w(\omega)H_y(\omega)H_y^*(\omega) = S_w(\omega)\left|H_y(\omega)\right|^2 \tag{8.15}$$

where $S_w(\omega)$ is the wave spectrum, $H_y(\omega)$ is the frequency response function, and * indicates the complex conjugate. Assuming structure reacts to wave excitation in a quasi-static way we have

$$H_y(\omega) = \sum_i c_i H_{f_i}(\omega) \tag{8.16}$$

where $H_{f_i}(\omega)$ is frequency response function of the *i*th wave induced load component, c_i is structure influence coefficient usually calculated by finite element method (FEM). Substituting Eq. (16) in (15) yields

$$S_y(\omega) = S_w(\omega)\left\{\sum_i c_i^2 f_i^2 + \sum_j \sum_k c_j c_k f_j f_k \cos(\alpha_j - \alpha_k)\right\} \tag{8.17}$$

where f is the amplitude of unit wave induced loads, and α is the phase angle.

8.3.2 Equivalent stress RAO based on pseudo-excitation

It should be noted that Eq. (17) is only applicable for linear systems. For non-linear loads a linear equivalent f has to be found. This can be achieved by evaluating the non-linear loads from a statistical wave height H_e, namely

$$f_i = \frac{\Delta F_i}{H_e} \tag{8.18}$$

where ΔF_i is the range of the *ith* load due to wave height H_e. Time domain simulation is needed to calculate ΔF_i in one wave cycle. Violette (1997) suggested significant wave height be used for H_e. Likewise a wave height corresponding to certain long term exceeding probability of stress range, for example 10^{-4}, is recommended by DNV (2003). Both approaches assume that intermediate wave amplitudes contribute most to the fatigue damage.

When dealing with random vibration problems, Lin (1992) looked at the spectral relationship in Eq. (15) from a new perspective. That is if we apply an imaginary excitation

$$x(t) = \sqrt{S_w(\omega)}e^{i\omega t} \tag{8.19}$$

, which gives stress response

$$y(t) = H_\sigma(\omega)\sqrt{S_w(\omega)}e^{i\omega t} \tag{8.20}$$

then the stress spectrum can be obtained readily by

$$S_y(\omega) = yy^* = |y|^2 \tag{8.21}$$

The excitation in Eq. (19) is called pseudo excitation. It reduces traditional random vibration analysis of linear systems into a series of deterministic harmonic analyses. If we extend this idea to Eq. (18) and let

$$H_e = 2\sqrt{S_w(\omega)} \tag{8.22}$$

the linearization can now be related to wave spectrum, which seems more appropriate for analysis in frequency domain. Since wave induced loads are in essence hydro responses from the same wave excitation, applying pseudo excitation defined in Eq. (19) gives

$$\boldsymbol{F} = \sqrt{S_w(\omega)}\begin{Bmatrix} H_{f_1}(\omega) \\ H_{f_2}(\omega) \\ \vdots \\ H_{f_n}(\omega) \end{Bmatrix} = \sqrt{S_w(\omega)}\begin{Bmatrix} f_1 e^{-i\alpha_1} \\ f_2 e^{-i\alpha_2} \\ \vdots \\ f_n e^{-i\alpha_n} \end{Bmatrix} e^{i\omega t} \tag{8.23}$$

Here linearised transfer functions should be used for non-linear loads. Accordingly the spectral matrix of the load components becomes

$$\begin{aligned} \boldsymbol{S}(\omega) &= \boldsymbol{F}^* \boldsymbol{F}^T \\ &= S_w(\omega)\begin{bmatrix} |H_{f_1}(\omega)|^2 & H^*_{f_1}(\omega)H_{f_2}(\omega) & \cdots & H^*_{f_1}(\omega)H_{f_n}(\omega) \\ H^*_{f_2}(\omega)H_{f_1}(\omega) & |H_{f_2}(\omega)|^2 & \cdots & H^*_{f_2}(\omega)H_{f_n}(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ H^*_{f_n}(\omega)H_{f_1}(\omega) & H^*_{f_n}(\omega)H_{f_2}(\omega) & \cdots & |H_{f_n}(\omega)|^2 \end{bmatrix} \end{aligned} \tag{8.24}$$

And we now have the following coherency function:

$$\frac{|\boldsymbol{S}_{ij}|^2}{\boldsymbol{S}_{ii}\boldsymbol{S}_{jj}} = 1 \tag{8.25}$$

This means the wave induced loads are fully coherent.

The wave load components in vector $\boldsymbol{F}$ can be applied simultaneously to a FE model to get the stress amplitude vector $\boldsymbol{\sigma}$ at hot spots. The spectral matrix of stress response can be obtained by

$$\boldsymbol{S}_y(\omega) = \boldsymbol{\sigma}^* \boldsymbol{\sigma}^T \tag{8.26}$$

whose diagonal components are equivalent to Eq. (17). This falls in direct load approach since load combination is automatically done for given loading condition, ship speed, ship heading and wave frequency. Typical number of load cases required in the FE analysis is given in Table 8.1.

Table 8.1 Number of load cases required in direct load approach.

Item	Number
Loading condition	2
Speed	5
Heading	19
Frequency	25
Sampling points in one wave cycle	10
Total number of load cases	47500

Monte Carlo simulation is useful to verify the narrow-band and Gaussian assumption. For coherent system as discussed above. According to Eq. (2.18) a FFT realization of the random processes can be generated as follows

$$\begin{aligned} f_i(t) &= \sqrt{2}\sum_0^{N-1} A_n \cos(\omega_n t + \Phi_n + \alpha_i) \\ &= \sqrt{2}\sum_0^{N-1} \sqrt{2S_{ii}(\omega_n)\Delta\omega}\cos(\omega_n t + \Phi_n + \alpha_i) \quad N \to \infty \end{aligned} \tag{8.27}$$

8.4 Wave induced responses

8.4.1 Types of cyclic loads

Cyclic loads experienced by a ship fall into four categories: still water loads, thermal loads, and low and high frequency wave induced loads. Static loads and thermal loads will affect the mean stress level.

Still water loads affects the mean values upon which wave load is superimposed. These loads do vary slightly within one voyage or may vary considerably between voyages. However this kind of cyclic behaviour does not contribute much to the total fatigue damage. This said, if a hull structure is subject to much higher frequency of loading and unloading as for FPSO variation of still water loads together with cargo loads must be considered in low cycle fatigue calculation.

Thermal stresses can be due to seasonal temperature changes or diurnal changes. The latter, though of higher frequency than the former, tends to correlate with calm sea states. So overall their cyclic contribution to the total damage is also marginal. Their effect is mainly associated with mean stress.

High frequency wave loads consist of transient and steady state ones. Transient loads are mainly due to slamming and whipping, steady state loads are usually related to springing. Slamming and whipping mainly happens to head sea and in extreme sea states, namely when the frequency of excitation is close to that of vertical motion. Operational measures are usually taken to avoid this. So their probability in the service profile is relatively low. Springing occurs in low to medium sea-states. The response is to do with hull girder natural frequency and hydrodynamic damping. It is not considered as major problem for ocean going ships as opposed to Great Lake ships with weaker section modulus.

As discussed above only low frequency wave induced motion, pressure distribution, hull girder loads, and cargo inertial loads are usually considered in spectral fatigue models.

8.4.2 Direct calculation procedures

Here we assume that the ship motions consist of small oscillations about an equilibrium position and that the properties of the fluid are determined by an irrotational, incompressible and in-viscid fluid flow theory. The total velocity potential around the ship consists of the steady part due to forward speed of the vessel and unsteady part due to ship oscillation in wave:

$$\Phi_T(x,y,z) = \Phi_S(x,y,z) + \Phi_W(x,y,z)e^{i\omega t} \tag{8.28}$$

The steady velocity potential can be written as

$$\Phi_S(x,y,z) = -Ux + \phi_S(x,y,z) \tag{8.29}$$

where U is the ship speed and ϕ_S is the steady disturbance potential. The unsteady potential can be decomposed as:

$$\Phi_W(x,y,z) = \phi_I + \phi_D + \sum_{j=1}^{6} \xi_j \phi_j \tag{8.30}$$

with

ϕ_I = incident wave potential

ϕ_D = diffracted wave potential

$\xi_j = j^{th}$ motion in six degree of freedom (Fig. 8.1)

ϕ_j = potential due to forced motion in j^{th} mode

ω = circular encounter wave frequency

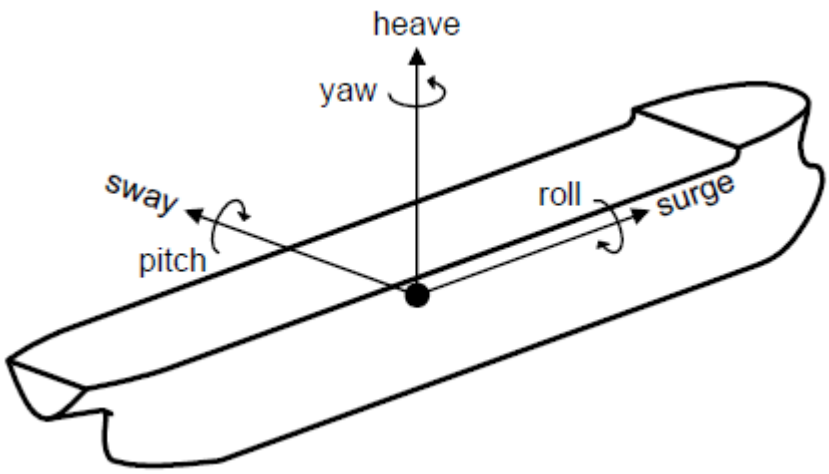


Figure 8.1 Ship motions in wave

The following boundary conditions must be satisfied

1) Laplace equation in the fluid domain:

$$\nabla^2 \Phi = 0 \tag{8.31}$$

2) Linearised free surface condition at $z = 0$:

$$\begin{aligned} &U^2 \frac{\partial^2 \phi_S}{\partial x^2} + g \frac{\partial \phi_S}{\partial z} = 0 \\ &\left[\left(i\omega - U \frac{\partial}{\partial x} \right)^2 + g \frac{\partial}{\partial z} \right] \left(\phi_I, \phi_D, \phi_j \right) = 0 \end{aligned} \tag{8.32}$$

3) Hull boundary conditions

$$\begin{aligned}
&\frac{\partial}{\partial n}\left(-Ux+\phi_S\right)=0\\
&\frac{\partial}{\partial n}\left(\phi_I+\phi_D\right)=0\\
&\frac{\partial \phi_j}{\partial n}=i\omega n_j+Um_j
\end{aligned} \tag{8.33}$$

where

$\left(n_1,n_2,n_3\right)=\boldsymbol{n}$, outward unit vector normal to the hull surface
$\left(n_4,n_5,n_6\right)=\boldsymbol{r}\times\boldsymbol{n}$, $\boldsymbol{r}$ is a position vector with respect to the origin

and

$$\left(m_1,m_2,m_3\right)=-\left(\boldsymbol{n}\cdot\nabla\right)\nabla\left(\phi_S-x\right)$$
$$\left(m_4,m_5,m_6\right)=-\left(\boldsymbol{n}\cdot\nabla\right)\left[\boldsymbol{r}\times\nabla\left(\phi_S-x\right)\right]$$

The pressure on the hull surface can be calculated by Bernouilli equation:

$$P=-\rho\left(\frac{\partial\Phi}{\partial t}+\frac{1}{2}\left|\nabla\Phi\right|^2+gz\right) \tag{8.34}$$

Usually only the linearised form of Eq. (34) is used by ignoring the higher order terms and terms involving cross products of steady and unsteady potential in the time dependent terms:

$$P=-\rho\left(i\omega-U\frac{\partial}{\partial x}\right)\phi e^{i\omega t} \tag{8.35}$$

The linear equation of motions can be written as:

$$\sum_{k=1}^{6}\left[-\omega^2\left(M_{jk}+A_{jk}\right)+i\omega B_{jk}\_C_{jk}\right]\xi_k = F_j, j=1,2,\ldots 6 \tag{8.36}$$

where the added mass and the damping coefficients are derived in terms of velocity potential by integrating the hydrodynamic pressure over the hull surface:

$$\omega^2 A_{jk} - i\omega B_{jk} = -\rho\iint n_j\left(i\omega - U\frac{\partial}{\partial x}\right)\phi_k dS \tag{8.37}$$

The wave excitation force in Eq. (35) can be expressed as:

$$F_j = -\rho\iint n_j\left(i\omega - U\frac{\partial}{\partial x}\right)\left(\phi_I+\phi_D\right)dS, j=1,2,\ldots 6 \tag{8.38}$$

The shear force and bending moments of the hull section at location x' can be written as:

$$\begin{aligned} Q(x') &= \int_{x'}^{bow}\left\{\frac{\mu(x)}{g}\left(\ddot{z}-x\ddot{\theta}\right)-F(x)\right\}dx \\ M(x') &= \int_{x'}^{bow}(x-x')\left\{\frac{\mu(x)}{g}\left(\ddot{z}-x\ddot{\theta}\right)-F(x)\right\}dx \end{aligned} \tag{8.39}$$

The velocity potential problem can be solved by 2D strip theory for slender ships (Korvin-Froukovski and Jacob 1957, Newman 1970). It implies that variation of the flow is much larger in transverse section than in longitudinal direction and the oscillation frequency is high. The 3D Laplace equation and the boundary conditions will be reduced to 2D. More importantly, the free surface condition in Eq. (32) will be reduced to one equivalent to that at the zero speed. Strip theory is relatively easy to implement and has been widely used. However under high frequency assumption it is more applicable for head sea than for following and quartering sea for a ship with forward speed. For

predicting motion it gives pretty good estimate compared to 3D theory even for ships with low length to beam ratio.

Wehausen and Laitone (1960) proposed Green's function that satisfies the linearised free surface condition. The nature of the Green's function is similar to Kelvin singularity and approaches the oscillating wave source in the limiting case of zero forward speed. Brard (1972) then found a way to express the radiation and diffraction potential in terms of Green's function. Chang (1977) compared Green's function approach with 2D strip theory and measured data for a series 60 hull. In general 3D results showed improvement over 2D ones but at much higher computing cost. Due to presence of singularity the Green's function with forward speed is difficult to integrate numerically. One approximation is to use a Green's function for zero speed (also by Wehausen and Laitone 1960). The forward speed is only considered through linearised pressure. In this research PRECAL a hydrodynamic software tool will be used to get the wave induced responses required for fatigue damage calculation (MARIN 2003).

8.5 Voyage simulation

A realistic long term wave environment is essential in determining the service profile matrix in Eq. (14). Much light has been shed on this in a study by Soares & Moan (1991). Voyage simulation procedure and software have been developed by Lloyd's Register for this purpose. It features 100A1 Fatigue Wave Environment that is comprised of a collective of statistical trading patterns for each ship type (Lloyd's Register 2004). On top of this user defined trading patterns can also be added. The simulation process can be summarized as follows:

- Exposure time to each sea state is derived from the maximum possible speed at the sea state and distance between the waypoints.

- The ship to wave headings are determined by the sailing course between waypoints and the wave direction in the global wave statistics (BMT 1986).

- Seakeeping criteria are defined to reflect the operational decisions at sea to avoid severe response.

- Parametric ship motion is calculated during simulation. Ship speed and heading are adjusted if necessary to meet the seakeeping criteria.

- Summate probabilities of occurrence for each combination of loading condition, ship speed, heading, significant wave height and average wave period of zero-upcrossings.

Operational measures will be taken to avoid occurrence of slamming, deck, wetness, excessive motions and accelerations. The following sea-keeping criteria are introduced in the simulation:

Table 8.2 Sea-keeping criteria in voyage simulation

Criteria	**Upper Limit**
Bow vertical acceleration	1.25g
Bridge vertical acceleration at 175% depth and 60% half beam	2g
Bridge lateral acceleration at 175% depth	2g
RMS (root of mean square) roll angle	6 degrees
Bow bottom slamming at 95% Lpp	3% occurrence
Deck wetness forward at 105% depth	10% occurrence
Propeller emergence at 10% depth	10% occurrence

Once any of the sea-keeping criteria is breached the speed is first reduced incrementally by 25% of the service speed until that criterion is met. If after speed has been reduced to zero and the response is still excessive the ship will be placed in the next adjacent heading. Considering the fatigue damage is mainly contributed by low to medium sea states it is believed that the total effect of sea-keeping criteria on calculated fatigue life is not significant. Voyage simulator implemented in ShipRight FDA2 software by Lloyd's Register (2002) will be used in the cases study.

8.6 FEM model

8.6.1Model extent and mesh size

For fatigue analysis in the cargo region it is enough to use three cargo hold model with hot spot located in the mid cargo hold. The mesh size in the hot spot area will be t by t in order to reflect stress concentration. The mesh size for the rest of the model may follow stiffer spacing. Since in spectral fatigue analysis we are only interested in the stresses in the local hot spot area the mesh size of coarse mesh in longitudinal direction can be up to half frame spacing. The fine mesh area must have enough extent and there must be smooth transition between coarse mesh and fine mesh area to minimize the effect of displacement boundary condition caused by such an idealization.

Alternatively stress analysis can be done in a top-down manner (also called global-local or zoom-in analysis). In this approach global coarse mesh and local fine mesh model are created separately. The global model will be loaded and analysed first. The displacements from the global model will then be transferred to the boundaries of local model along with the internal loads in the fine mesh area. The local model is usually of much smaller scale and can be run more quickly in the subsequent analyses. This approach is suitable when the computing resource is a limiting factor or when different teams need to collaborate in global and local modelling. However a single model with embedded fine mesh is preferred in the current study as it allows more than one local fine mesh area to be analysed in one go. Besides, modern PCs can deal with FE model of such a scale easily.

8.6.2 Boundary conditions

Planar boundary condition will be applied to the ends of the three cargo hold model. This is achieved though multi-point constraint (MPC) with master point located at the neutral axis height and slave points at the ends of the continuous longitudinal structures (Fig. 8.2). MPC can be used to link the master and slave points at d.o.f (degree of freedom) 2, 5, 6 or 1, 2, 3. If we view hill girder as a beam the former corresponds to Euler theory and

the latter Timoshenko theory. Both treatments allow bending moments defined at the master point to be applied to the entire transverse section at the same location. To remove rigid body motion single point constraint is applied to the forward master point at d.o.f 1, 2, 3, 4 and after master point at d.o.f 2, 3, 4.

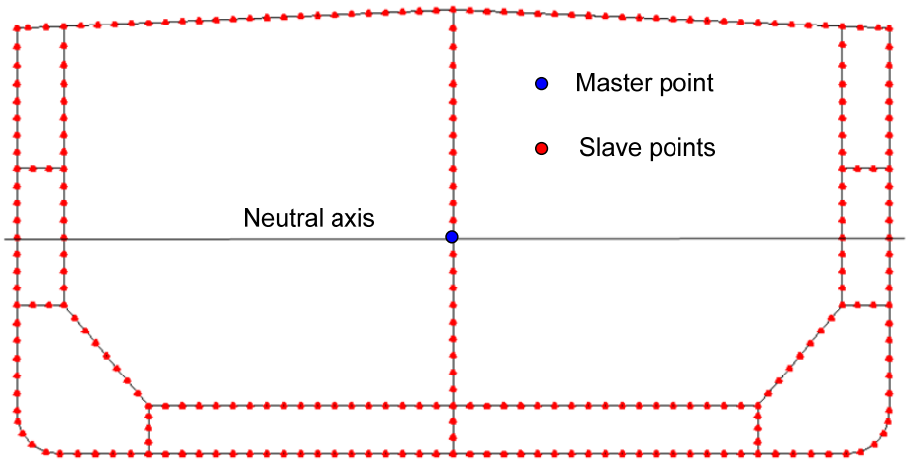


Figure 8.2(a) MPC constraints for double hull oil tanker

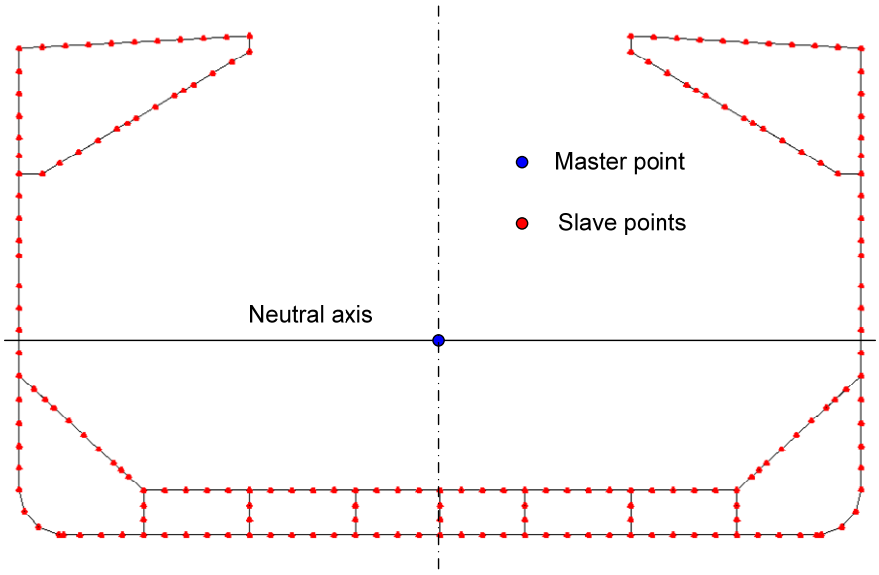


Figure 8.2(b) MPC constraints for single hull bulk carrier

8.6.3 Unit load cases

Total number of combination in direct load approach can make even quasi-static analysis a daunting task. To speed up calculation, a set of predefined unit load cases can be used. It consists of unit section loads, unit patch loads and unit inertia loads. Stresses obtained from unit load cases are the structural influence coefficients in Eq. (16), which are used to relate stress amplitude to wave amplitude for different operational and sea states.

In a three-cargo-hold model unit section loads are bending moments and shear forces applied to the model ends. To transfer pressure from hydrodynamic analysis to the FE mode patches are defined over the hull form. Each patch is formed by a group of adjacent elements in the FE model within which the wave pressure is assumed to be a constant. Unit pressure is then applied to each patch. In parallel part of the hull form parches are often quadrilateral areas. Apparently patch defined as such will be coarser than hydro mesh. Its size should be decided by the variation of hydrodynamic pressure (Lloyd's Register 2004). Pressure at patch centroid can be obtained by bilinear interpolation from pressures at four surrounding facet centroids in the hydro mesh (Fig. 8.3).

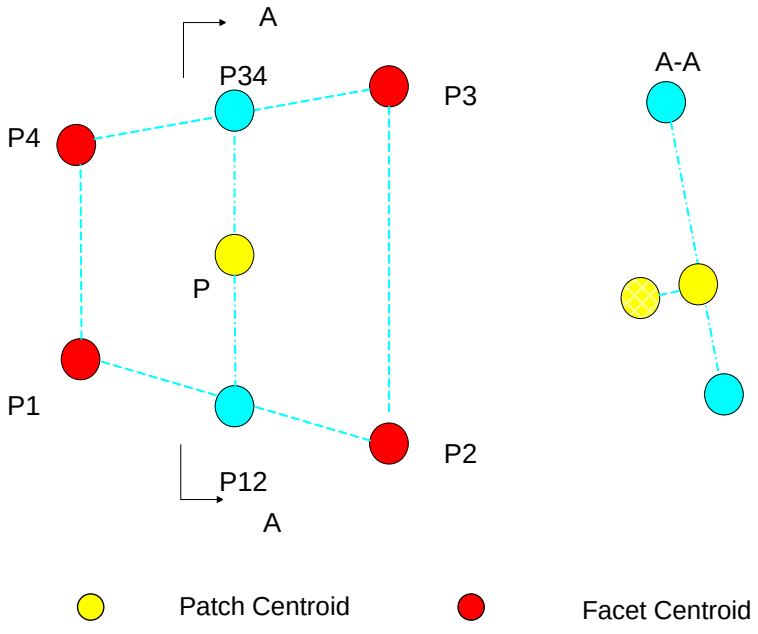


Figure 8.3 Pressure transfer through bilinear interpolation

Where such an interpolation is impossible the pressure of the closest hydro mesh may be used. Another useful approximation is the use of the weighted average:

$$P_{patch} = \frac{P_1 d_2 d_3 d_4 + P_2 d_1 d_3 d_4 + P_3 d_1 d_2 d_4 + P_4 d_1 d_2 d_3}{d_2 d_3 d_4 + d_1 d_3 d_4 + d_1 d_2 d_4 + d_1 d_2 d_3} \tag{8.40}$$

where P_1 to P_4 are the pressures at the four closest hydro facets to the patch under concern, d_1 to d_4 are the distances between patch and facet centroids. An example of hydro facet, patch and FE mesh is given in Fig. 8.4.

Linear inertia pressure induced by unit translational accelerations is illustrated in Fig. 8.5. The zero reference point is taken at the tank centroid (Violette 1997) for liquid cargo holds. Gravity components due to roll and pitch can be considered through unit load cases defined in Fig. 8.6. The zero reference point is the highest point in the tank. These are nonlinear load cases by nature.

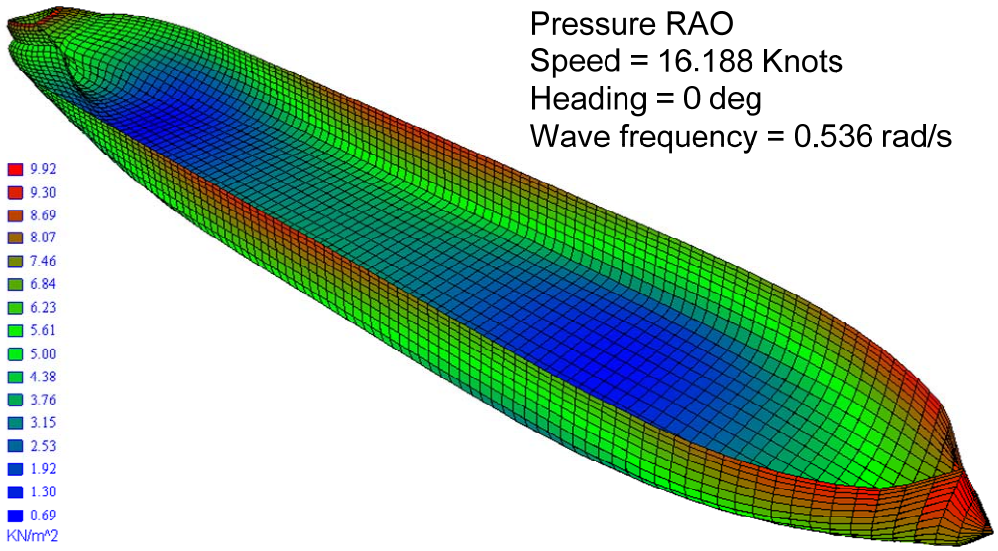


Figure 8.4(a) Example of hydro mesh with pressure in head sea

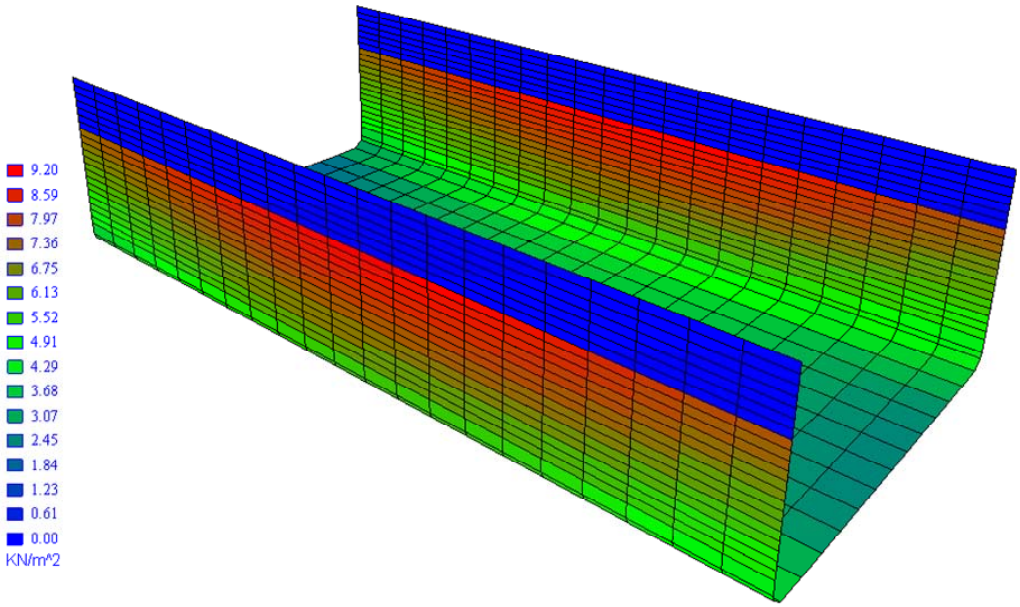


Figure 8.4(b) Transferred hydro pressure on patch

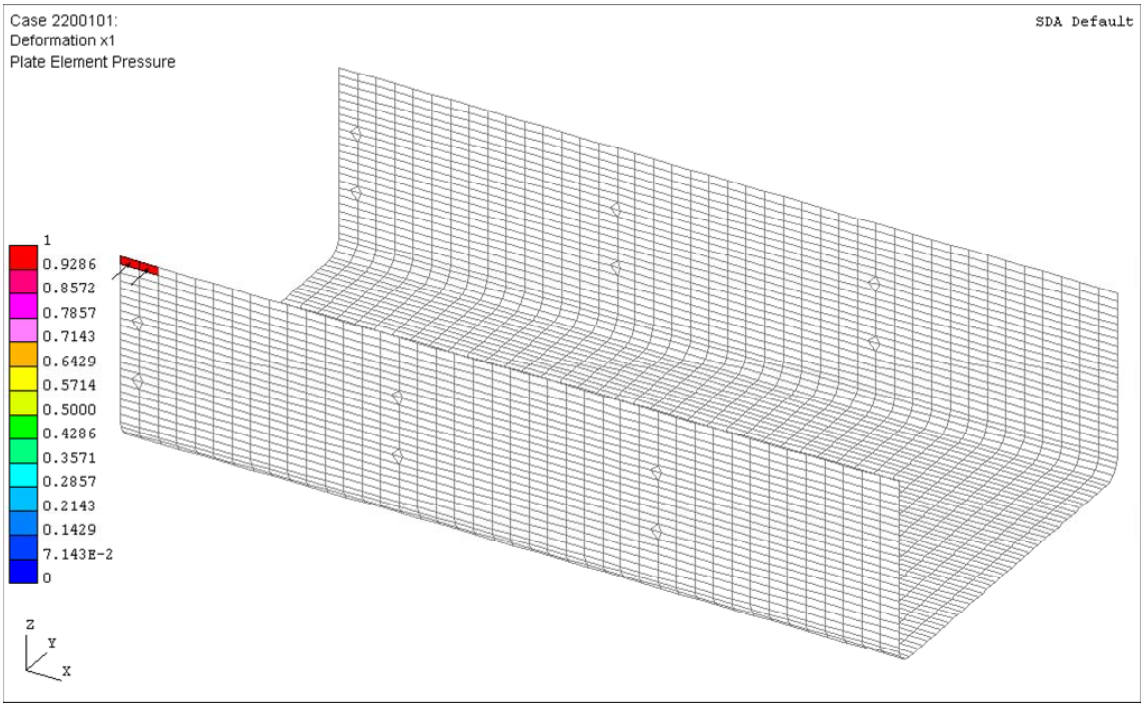


Figure 8.4(c) FE mesh for side shell with unit patch load

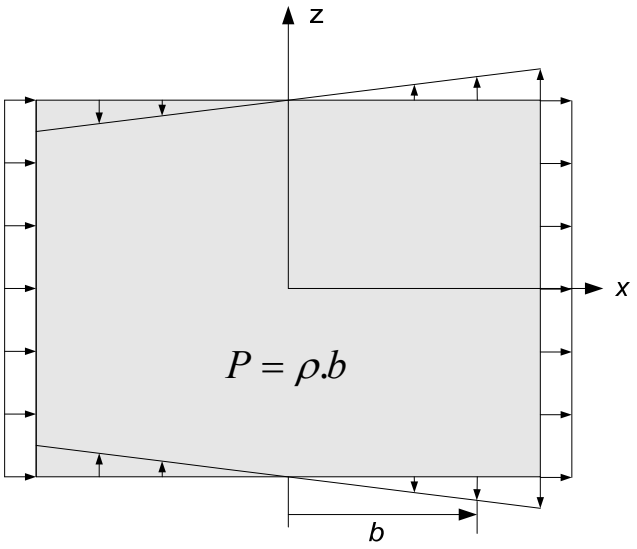


Figure 8.5(a) Pressure due to longitudinal acceleration

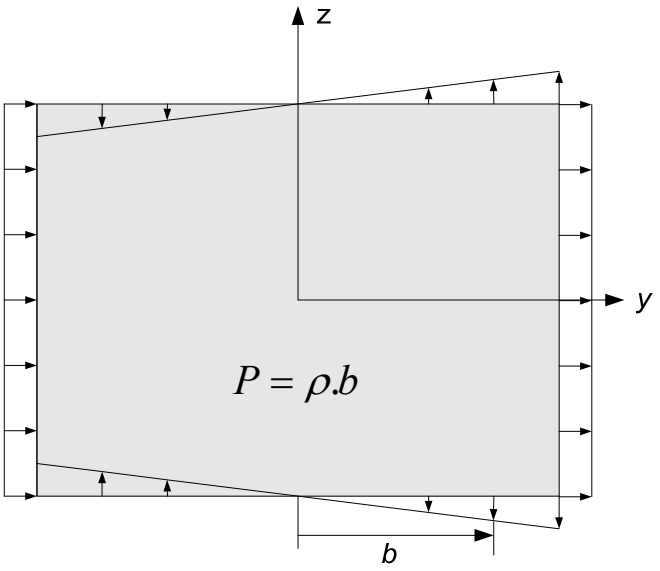


Figure 8.5(b) Pressure due to transverse acceleration

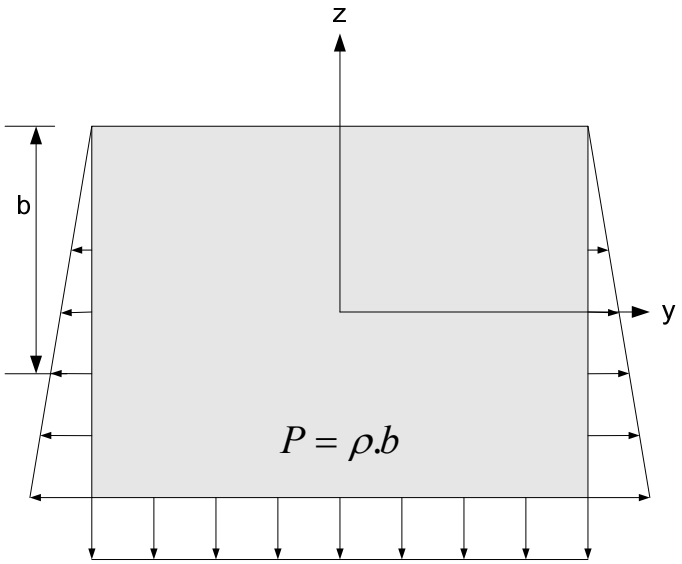


Figure 8.5(c) Pressure due to vertical acceleration

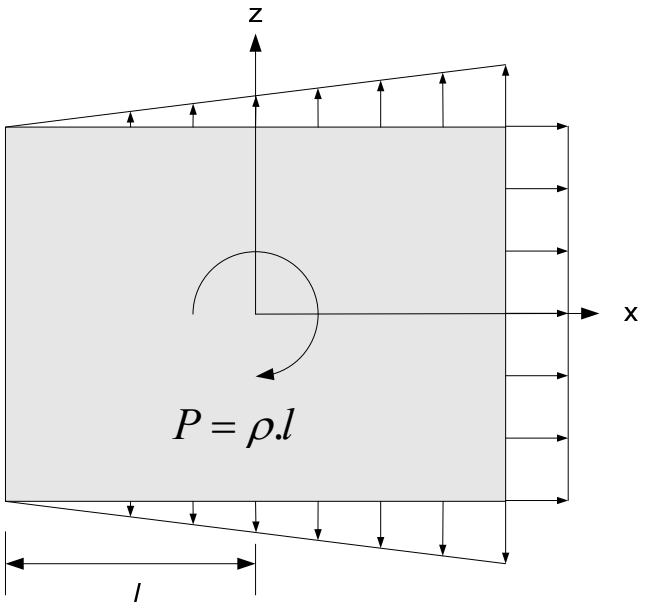


Figure 8.6(a) Pressure due to positive pitch

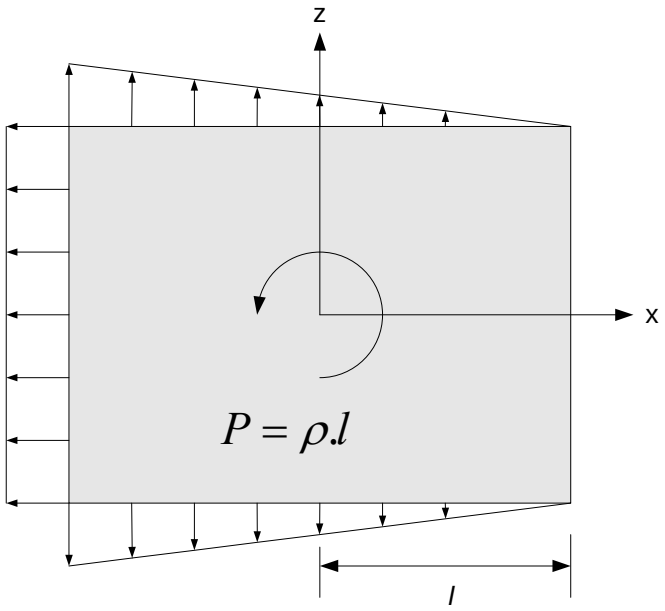


Figure 8.6(b) Pressure due to negative pitch

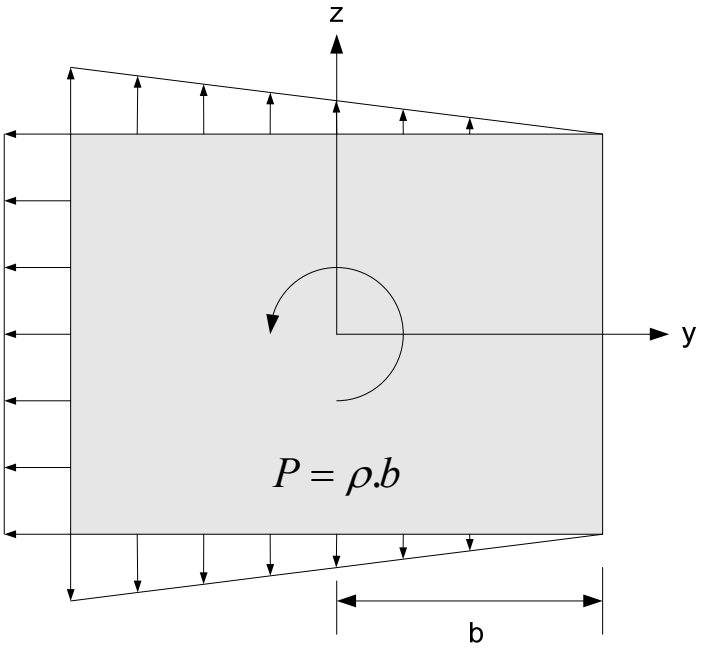


Figure 8.6(c) Pressure due to positive roll

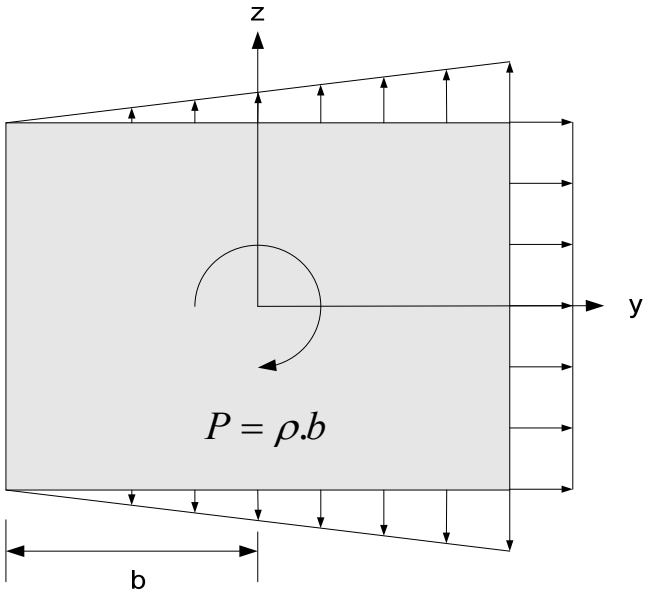


Figure 8.6(d) Pressure due to negative roll

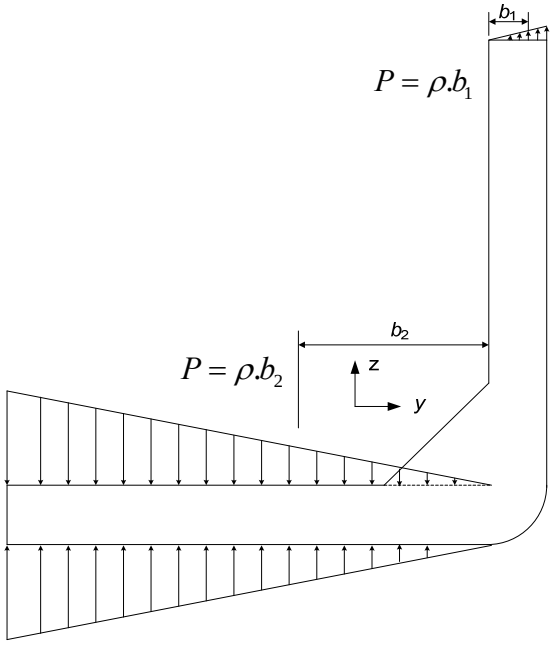


Figure 8.6(e) Pressure due to negative roll for L shaped tank on the port side

8.7 Case study

8.7.1 Double hull tanker

8.7.1.1 Deterministic model for as-designed connection

This is a damage case for a typical Aframax double hull oil tanker. The hot spot locates at the intersection between side stringer, inner hull, transverse bulkhead and the second horizontal stringer attached to it (Fig. 8.7). Two loading conditions are considered: ballast and full load. The structural detail is pretty close to the draft in the full load condition. So the non-linearity in the intermittent wetting area is expected to have significant effect on the fatigue damage.

PRECAL, a linear seakeeping program based on 3D potential theory, is used to calculate the hydro responses. ShipRight SDA 2007, a FE based system developed by Lloyd's Register for SDA and FDA3, is used to create the unit fatigue load cases and boundary conditions.

ITTC wave spectrum is used in the calculation:

$$S(\omega) = 173 H_s^2 T_1^{-4} \omega^{-5} \exp(-691 (T_1 \omega)^{-4}) \tag{8.41}$$

where H_s is the significant wave height and T_1 is the characteristic period. Cosine-squared spreading function is applied to show the effect of directional wave:

$$D(\theta) = \begin{cases} \frac{2}{\pi} \cos^2(\theta - \theta_0), \text{for} \left(-\frac{\pi}{2} + \theta_0 \right) < \theta < \left(\frac{\pi}{2} + \theta_0 \right) \\ 0, \text{Otherwise} \end{cases} \tag{8.42}$$

where θ_0 is the mean wave direction.

For each element in the hot spot area fatigue damage is calculated for a range of fracture planes over ±90 degrees to the element local axis X (Fig. 8.8). The two slope SN curve

for fillet weld defined in FDA3 procedure (Lloyd's Register 2004) is used (Table 8.3). The fatigue life is shown in Fig. 8.9. The maximum fatigue damage is found at inner hull with the fracture plane being in vertical direction. The element is immediately after the transverse bulkhead and below the side stringer. It is also noticed that directional wave gives slightly longer fatigue life. The calculation shows the structure detail falls short of expected fatigue strength. Through thickness crack tends to develop during early years of service. This has been proved by data collected in ship survey.

Dirlik distribution is used for short term fatigue damage in each sea state. As shown in Table 8.4, it agrees very well with Wirsching's correction in this example. The long term distribution of stress range is given in Fig. 8.10.

In Miner-Palmgren linear accumulation theory long term fatigue damage is independent of loading history. So fatigue damage in different sea states can be calculated in parallel. It is very suitable for modern computers with multi-core processor. In the current example four computing threads are created during calculation through OpenMP, a multi-processing application program interface (API). It is found that the performance can be increased by 48% on a workstation with Intel Core Duo processor.

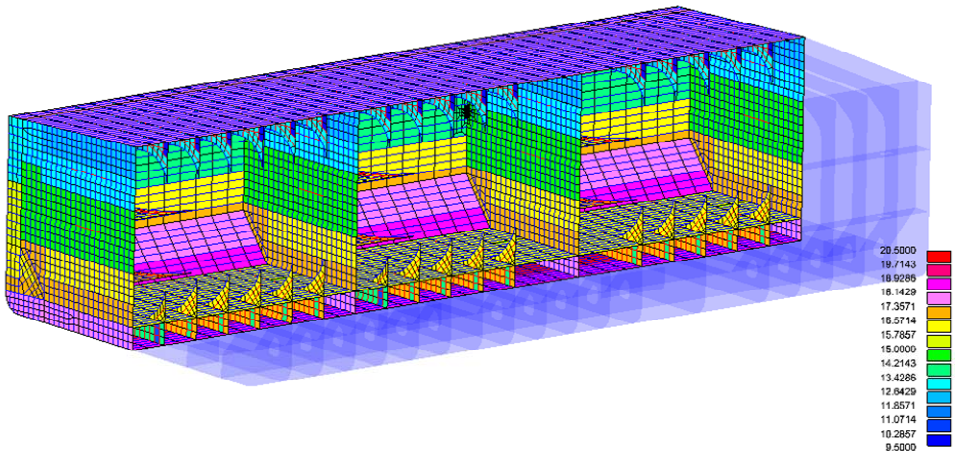


Figure 8.7(a) Three cargo hold model with embedded fine mesh

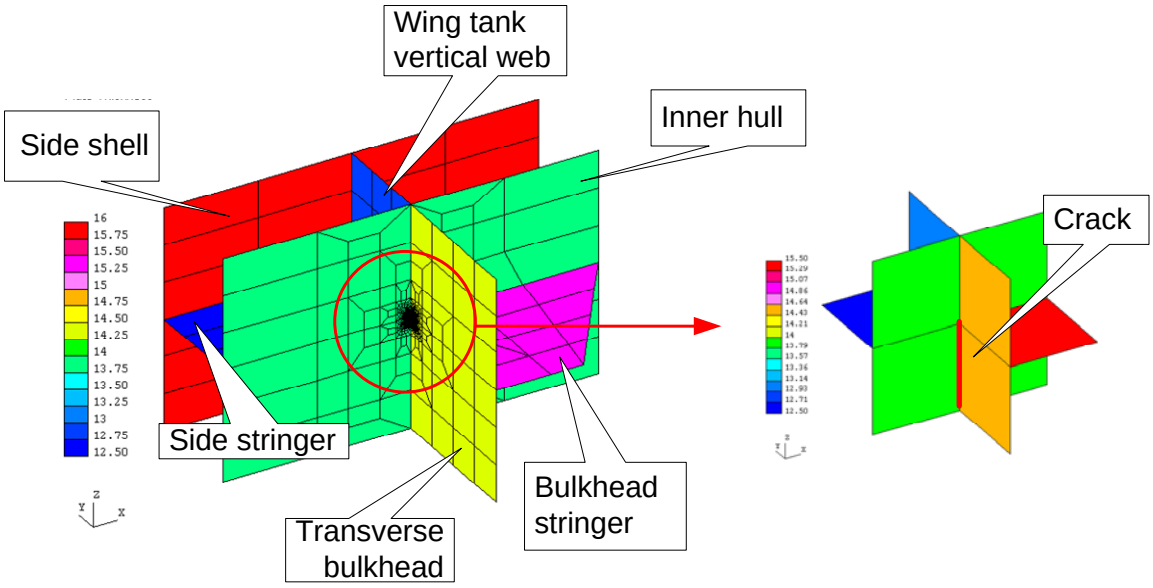


Figure 8.7(b) Local t by t fine mesh zone

Table 8.3 SN curve for fillet weld

Log10(K)	**m**	**m'**	**S_q**	**Standard Deviation**
12.636	3	5	75.625	0.2218

Table 8.4 Comparison of minimum fatigue life

Method	Mean Curve		Design Curve	
	L1	L2	L1	L2
Dirlik	8.413	10.081	2.962	3.521
Wirsching	8.514	10.106	2.999	3.531

L1 Fatigue life without spreading
L2 Fatigue life with spreading

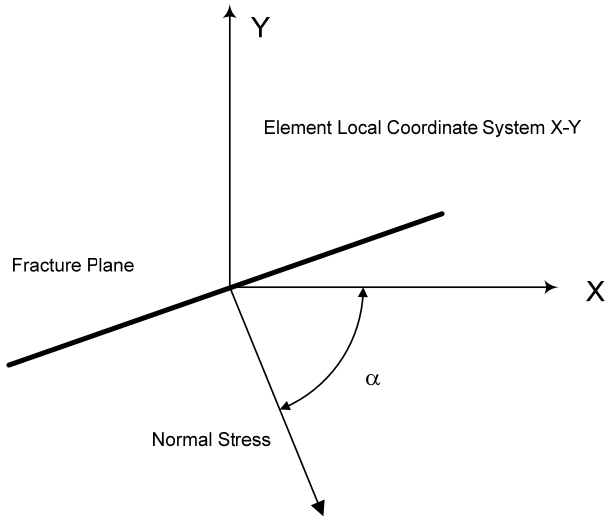


Figure 8.8 Definition of critical fracture plane

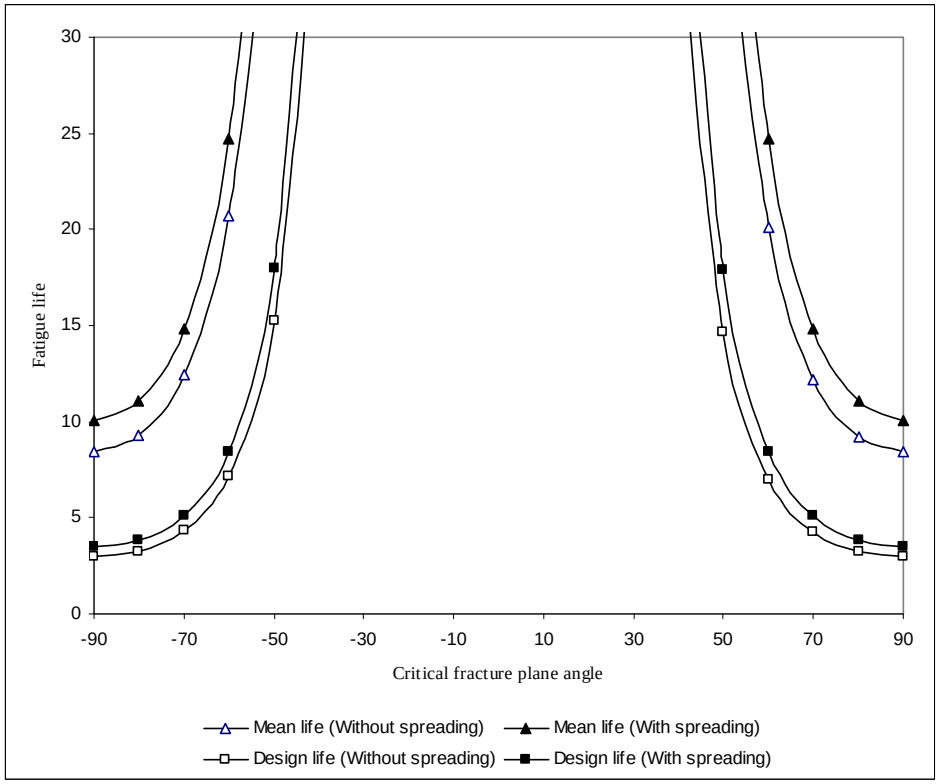


Figure 8.9 Fatigue of critical fracture planes

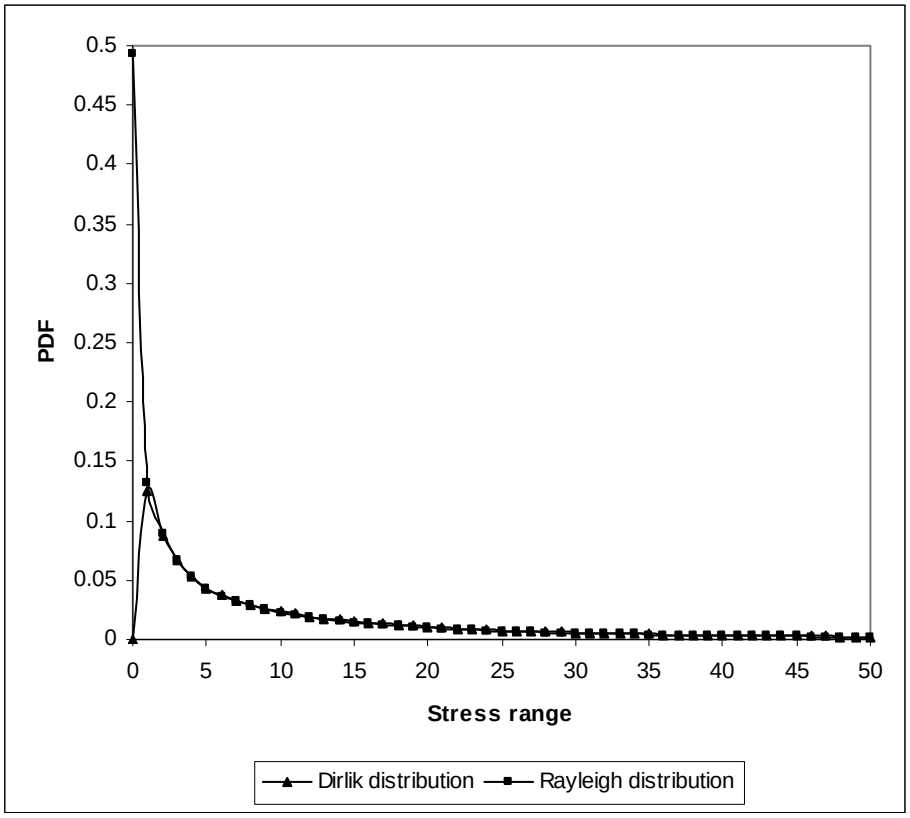


Figure 8.10 Long term distribution of stress range

8.7.1.2 Reliability model for enhanced connection

Since the original design is short of the required fatigue life a backing bracket is added to strengthen the interconnection (Fig. 8.11). Initial estimate shows the maximum fatigue damage is from side stringer at aft toe of the backing bracket. The design fatigue life is 49 years using SN curve in Table 8.3. Dirlik distribution and wave spreading are considered in the calculation. Define limit state function:

$$g(\boldsymbol{X}) = T - 20 \tag{8.43}$$

where $\boldsymbol{X}$ is the vector of design variables and T is the long term fatigue life. Random variables in the limit state function are given in Table 8.5.

Step-wise response surface method is then employed to fit a quadratic form of Eq. (43). In each adaptive step 21 samples are used, invoking 21 full FDA3 analyses. As shown in Fig. 8.12 the whole process shows steady convergence for the current problem. It only takes 7 steps to arrive at a RSF with relative error of 1.0e-3 compared to the last step. Two extra steps are run to further validate the converged point. Finally a relative error of 1.0e-4 is achieved at step 9 with the number of sample points totalling 189. The final design points and gradients are given in Table 8.6. Sensitivity of reliability with respect to design point in U space, mean value, and standard deviation is given in Table 8.7. It can be noticed that K, E, T_1, and T_2 have the most significant effects in the order of their importance. Thickness of side stringer T_4 is only in the fifth place. It is also interesting to see that increasing the thickness of backing bracket T_1 will actually lower the fatigue strength at current location. This is because without backing bracket the maximum fatigue damage would have happened right at the corner of intersection (i.e. the current bracket heel). The presence of the bracket simply transfers the risk away to its aft toe location.

A complete quadratic form for the current problem contains 65 non constant terms. Of these 34 are picked up in the final RSF by their contribution (Table 8.8). To study the non-linearity of the final RSF at design point SORM and MCS are apply to refine the calculated probability of failure. The comparison with FORM is given in Table 8.9. It indicates weak non-linearity of obtained RSF in the standardised space U.

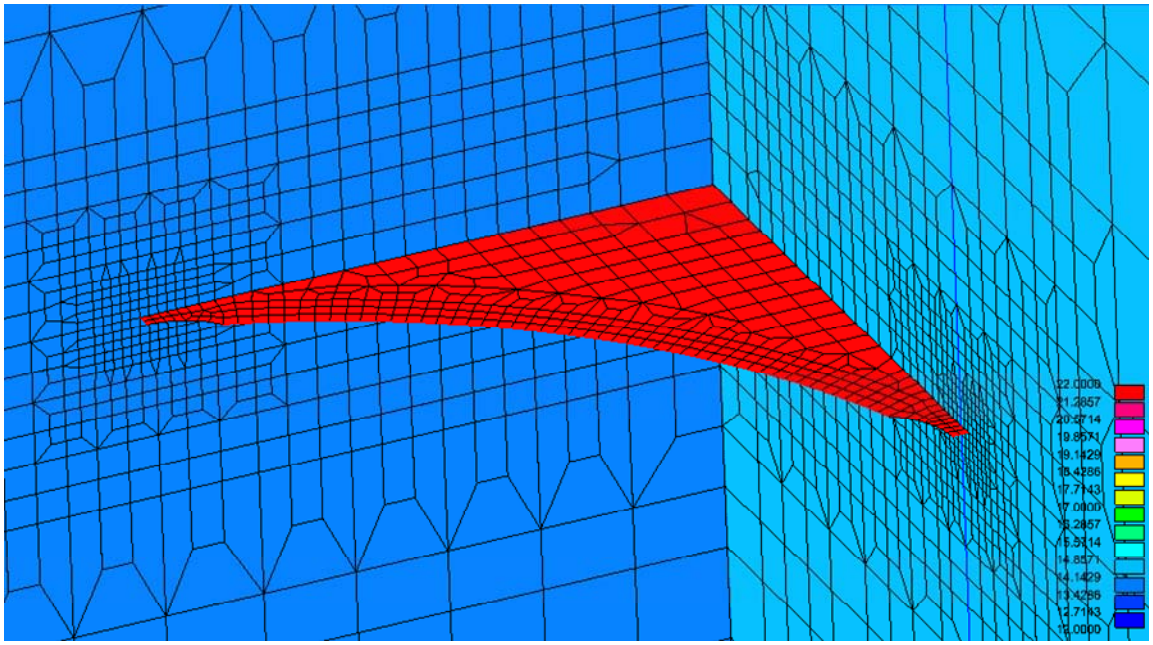


Figure 8.11 Enhanced design with backing bracket

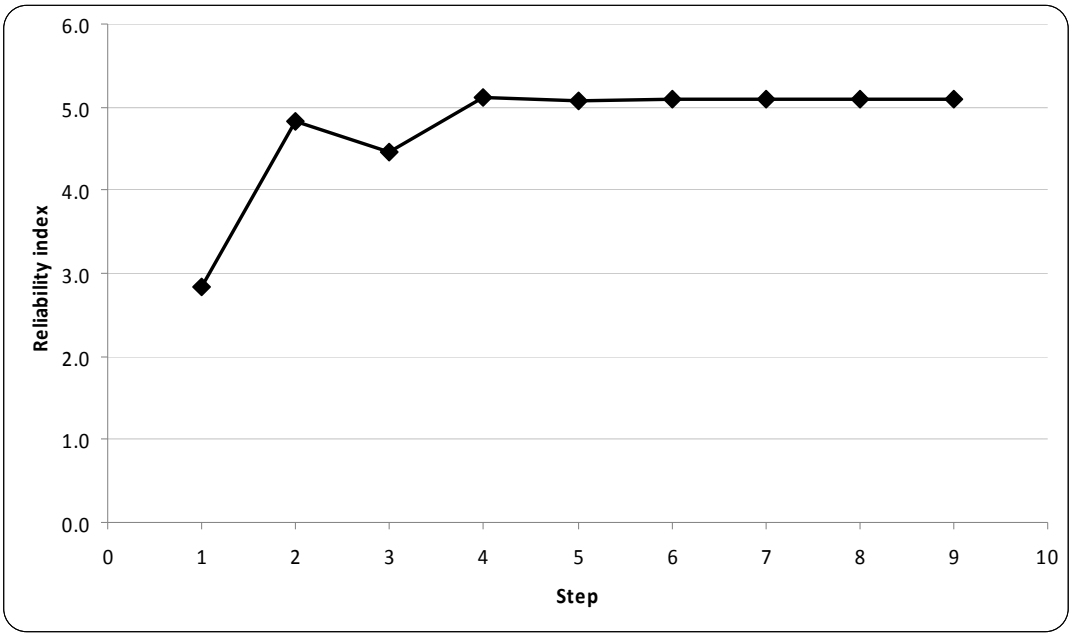


Figure 8.12 Reliability index in stepwise RSM iteration (Case study 1)

Table 8.5 Random variables (Case study 1)

Random Variable	Distribution	Mean Value	COV
Young's Modulus E	Lognormal	206GPa	0.07
Backing bracket thickness T_1	Normal	22 mm	0.02
Inner hull thickness T_2	Normal	14mm	0.02
Side shell thickness T_3	Normal	16mm	0.02
Side stringer thickness T_4	Normal	12mm	0.02
Transverse bulkhead horizontal stringer thickness T_5	Normal	15.5mm	0.02
Transverse bulkhead thickness T_6	Normal	14.5mm	0.02
Upper wing tank vertical web thickness T_7	Normal	15.0mm	0.02
Lower wing tank vertical web thickness T_8	Normal	13.0mm	0.02
Intercept of SN curve K	Lognormal	4.58E12	0.346

Table 8.6 Final design point and gradients (Case study 1)

RANDOM VARIABLE	$\boldsymbol{X}^*$	$G_{,\boldsymbol{X}^*}$	$G_{,\boldsymbol{U}^*}$
E	0.214E+06	-3.17E-4	-2.04
T_1	22.230	-1.77	-7.78E-1
T_2	13.876	2.36	6.62E-1
T_3	16.003	-5.19E-2	-1.66E-2
T_4	11.924	1.97	4.72E-1
T_5	15.529	-4.43E-1	-1.37E-1
T_6	14.505	-8.61E-2	-2.50E-2
T_7	14.999	1.92E-2	5.77E-3
T_8	13.002	-5.13E-2	-1.33E-2
K	0.846E+12	2.53E-11	7.22

Table 8.7 Sensitivity analysis (Case study 1)

RANDOM VARIABLE	$\beta_{,\boldsymbol{U}^*}$	$\beta_{,mean}$	$\beta_{,sd}$
E	2.69E-1	-4.18E-5	-5.83E-5
T_1	1.03E-1	-2.33E-1	-1.22E-1
T_2	-8.72E-2	3.12E-1	-1.38E-1
T_3	2.19E-3	-6.85E-3	-7.65E-5
T_4	-6.22E-2	2.59E-1	-8.21E-2
T_5	1.81E-2	-5.84E-2	-5.38E-3
T_6	3.29E-3	-1.13E-2	-1.90E-4
T_7	-7.61E-4	2.54E-3	-9.84E-6
T_8	1.76E-3	-6.76E-3	-6.06E-5
K	-9.51E-1	1.64E-12	-2.94E-12

Table 8.8 RSF coefficients (Case study 1)

RANDOM VARIABLE	**Coefficient**	**Partial Correlation** R_i
E	-3.3684E-03	-0.3525
T_1	1.6948E+01	0.1114
T_2	-2.2904E+01	-0.1638
T_3	-1.9509E+01	-0.1250
T_4	-3.1031E+01	-0.1015
T_5	-1.2018E+01	-0.1425
T_6	-9.0017E+00	-0.0980
T_7	-9.2660E+00	-0.1043
T_8	-2.6752E+01	-0.1323
K	1.0155E-10	0.8333
E^2	6.7522E-09	0.7519
T_1^2	3.8814E-01	0.2867
T_2^2	5.3065E-01	0.1621
T_3^2	2.7474E-01	0.1093
T_4^2	6.0937E-01	0.1303
T_5^2	3.9578E-01	0.1457
T_6^2	3.1082E-01	0.0982
T_7^2	3.0953E-01	0.1045
T_8^2	4.1858E-01	0.1094
K^2	1.4659E-24	0.9837
T_1E	-9.3978E-06	-0.0360
T_2T_1	-1.0976E+00	-0.2108
T_4E	4.7240E-05	0.0663
T_4T_1	-1.4081E+00	-0.1689
T_4T_2	2.6788E+00	0.1572
T_8E	1.3234E-05	0.0409
T_8T_3	8.2014E-01	0.1116
KE	-4.4039E-16	-0.9829
KT_1	-2.2908E-12	-0.8981
KT_2	3.5458E-12	0.8944
KT_4	2.9102E-12	0.7805
KT_5	-8.4750E-13	-0.4101
KT_6	-1.1953E-13	-0.0588
KT_8	-1.7578E-13	-0.0865
Regression Constant = 0.1052E4		

Table 8.9 Comparison with SORM and MCS (Case study 1)

	FORM	**SORM**	**MCS**
β	5.093	5.139	5.139
P_f	1.763E-7	1.383E-7	1.382E-7

SORM: Direct integration based on curvature fitting
MCS: 10^8 samples drawn at design point, cov = 2.49E-4

8.7.2 Bulk carrier

8.7.2.1 About the damage case

The sample ship is a Panamax size bulk carrier with 7 cargo holds. Cargo hold No.4 also doubles as the ballast hold, where the fatigue cracks were found during a hull survey 6 years into the ship's service. The cracks developed at inner bottom to lower stool connection in way of the double bottom girders as shown in Fig. 8.13. Only the location at the centreline girder is modelled and analysed in this study.

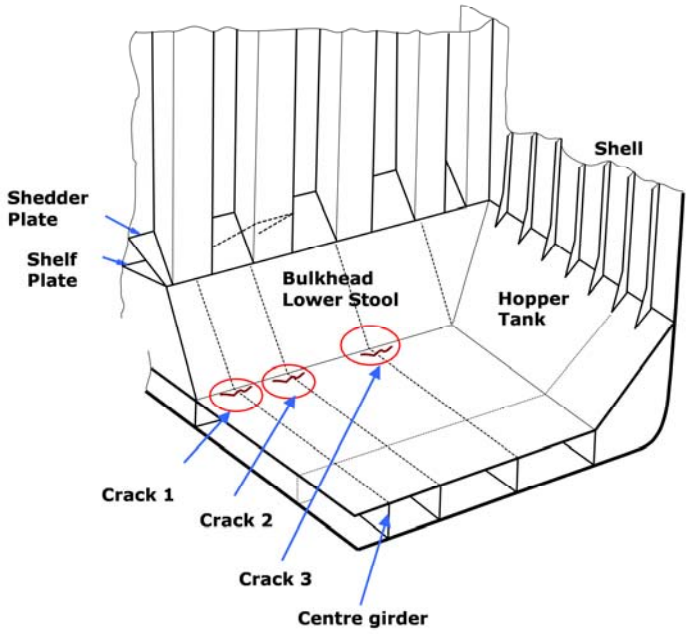


Figure 8.13 Locations of detected damage

The ship has been trading worldwide. The trading history of the ship was provided in the form of the ship's log containing local time of arrival and departure and cargo types as loaded or discharged. Previous study (Zhou et al. 2009) showed that of the possible ballast conditions the special light ballast condition gives maximum fatigue damage when combined with grain, coal and ore conditions (see Fig. 8.14). This combination will be studied in the reliability model with fraction given below:

Ballast loading condition : 25%
Iron Ore loading condition : 4.48%
Coal loading condition : 17.24%
Grain loading condition : 23.97%
Non-sailing : 29.31%

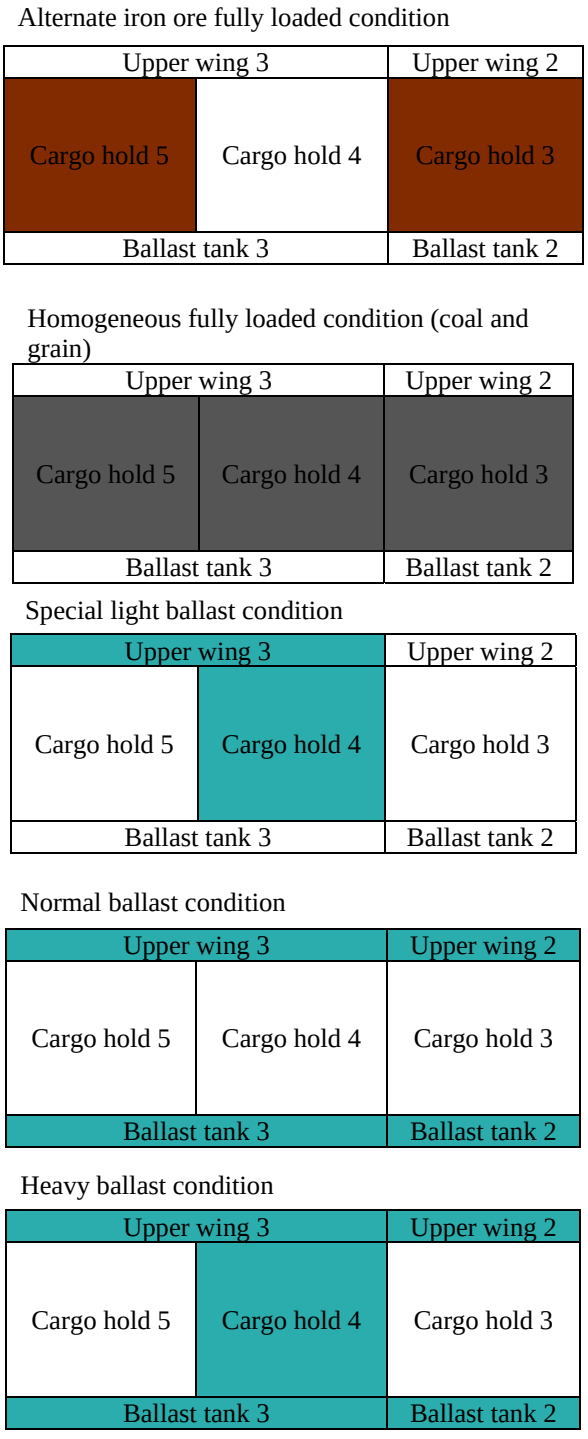


Figure 8.14 Arrangements of cargo and ballast loading conditions.

The global model consists of cargo holds No. 3 through No.5. Boundary condition and unit load cases are applied to the model in compliance with ShipRight FDA3 procedure. The number of load cases totals 1317, including 1216 patch loads, 97 inertial loads and 4 section loads. Stresses from the unit load cases will be scaled by corresponding hydrodynamic responses under unit wave excitation to get stress RAO. To speed up FE analyses in RSM a sub-model is created with a t by t fine mesh zone in the hotspot area. Boundary displacements and local loads are transferred from the global FE model, which only needs to be run once. Since the variations of the scantlings are small it is assumed that boundary displacements calculated at mean values of the design variables do not need to be updated in the subsequent RSM iterations. Accordingly the extent of the sub-model is carefully chosen to cater for this assumption and to include all the structure parts to be randomized. All stiffeners in the fine mesh region are represented using plate elements. This approach gives tenfold increase in speed. The only major trade-off is it tends to overestimate the importance of Young's modulus E. Since E is not a design factor to consider in practice the idealisation is believed acceptable. The FE model is given in Fig. 8.15.

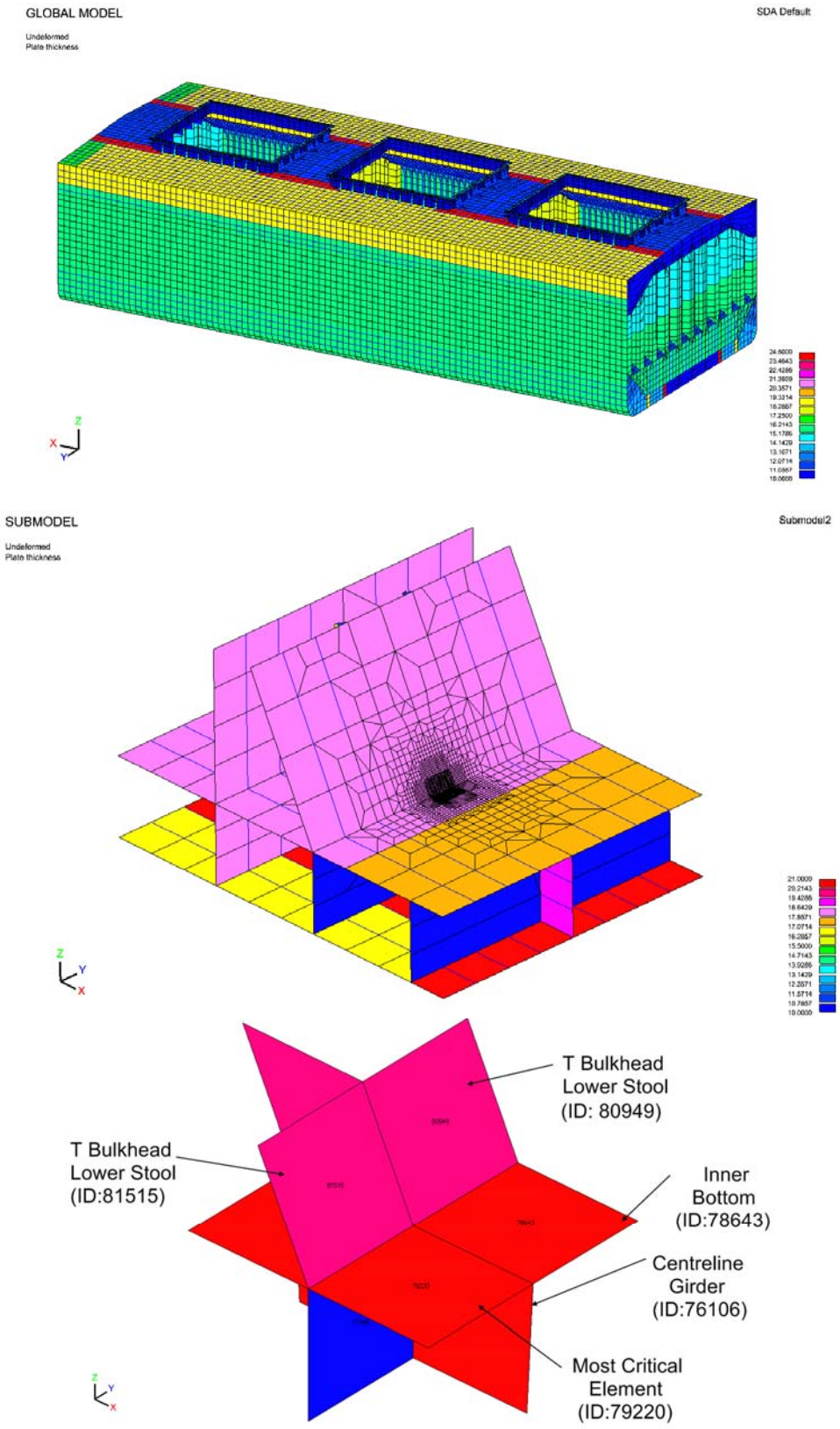


Figure 8.15 Global and local FE mesh

8.7.2.2 Reliability model

Standard FDA3 analysis shows the structural detail under concern only has fatigue life of 6.33 years using mean SN curve in Table 8.3. The maximum damage happens in the element immediately to the portside of the centreline girder with primary crack plane parallel to the knuckle line. This agrees well with the survey data but way below the expected service life.

If overhaul of the design is something formidable it might be worth aligning the repair with special surveys, which is due every 5 years. Probability of failure before the first special survey in service is looked at using limit state function:

$$g(\mathbf{X}) = T - 5 \tag{8.44}$$

where $\boldsymbol{X}$ is the vector of design variables and T is the fatigue life from the mean curve. Random variables in the limit state function are listed in Table 8.10.

A quadratic form of RSF is then obtained through stepwise RSM. In each adaptive step 27 samples are required, involving 25 FE analyses. The scantlings of the sub-model are updated on the fly. The whole process shows steady convergence for the current problem (Fig. 8.16). It can be noticed that the linear approximation in the first two steps gives better than expected estimate, implying simplicity of the actual response surface at design point. Once the second order terms are introduced in step 3 the reliability index is quickly brought within an absolute error of 5.5E-5 compared to the final value at step 7. The following 4 iterations only grind the error further down and finally result in reliability index β = 0.661, and probability of failure P_f = 0.254.

The final design points and gradients are given in Table 8.11. Sensitivities with respect to design point in the standard normal space U, mean values, and standard deviations are given in Table 8.12. It is clear that increasing K, T_1, and T_7 is the most effective way to improve the fatigue strength and the reliability level. It is also interesting to see that thickness increase in T_2, T_4, T_8 and T_{11} has reverse effect on fatigue life because this would impose more constraint on the hotspot elements. As we mentioned earlier the

importance of E is overestimated in the sensitivity result because of the displacement boundary condition in the sub-model. This is confirmed by rerunning the global FE model using variations of E.

A complete quadratic form for the current problem contains 104 non constant terms. Of these 34 are picked up in the final RSF, including all the linear and square terms and 8 cross terms (Table 8.13). To study the non-linearity of the final RSF at design point SORM and MCS are apply to refine the calculated probability of failure. The comparison with FORM is given in Table 8.14. It indicates weak non-linearity of the obtained RSF in the U space. To improve the fatigue strength T_7 is increased to 22mm and T_1 to 12.5mm. A second round of stepwise RSM is carried out, which gives $\beta = 1.989$, and $P_f = 0.0234$. (Fig. 8.16). SORM and MCS refinement is given in Table 8.15.

Table 8.10 Random variables (Case study 2)

Random Variable	Description	Distribution	Mean Value	COV
E	Young's Modulus E	Lognormal	207GPa	0.03
T_1	Double bottom floor	Normal	10 mm	0.02
T_2	Double bottom girder	Normal	19mm	0.02
T_3	Double bottom girder stiffener	Normal	12.5mm	0.02
T_4	Stool diaphragm	Normal	18mm	0.02
T_5	Inner bottom stiffener web	Normal	11mm	0.02
T_6	Inner bottom stiffener flange	Normal	16mm	0.02
T_7	Inner bottom plate strake 1	Normal	18.5mm	0.02
T_8	Inner bottom plate strake 2	Normal	17.5mm	0.02
T_9	Stool forward side plate stiffener web	Normal	11mm	0.02
T_{10}	Stool forward side plate stiffener flange	Normal	16mm	0.02
T_{11}	Stool forward side plate	Normal	18mm	0.02
K	Intercept of SN curve	Lognormal	4.58E12	0.346

Table 8.11 Final design point and gradients (Case study 2)

RANDOM VARIABLE	X^*	$G_{,X^*}$	$G_{,U^*}$
E	2.079E5	-7.238E-05	-4.514E-01
T_1	9.998	1.181E-01	2.362E-02
T_2	19.002	-3.846E-02	-1.461E-02
T_3	12.500	6.486E-03	1.622E-03
T_4	18.006	-1.171E-01	-4.215E-02
T_5	11.000	8.380E-03	1.844E-03
T_6	16.000	1.536E-03	4.915E-04
T_7	18.471	5.766E-01	2.133E-01
T_8	17.503	-7.405E-02	-2.592E-02
T_9	11.000	-2.867E-03	-6.307E-04
T_{10}	16.000	1.165E-05	3.729E-06
T_{11}	18.018	-3.819E-01	-1.375E-01
K	3.495E13	1.466E-12	1.726E+00

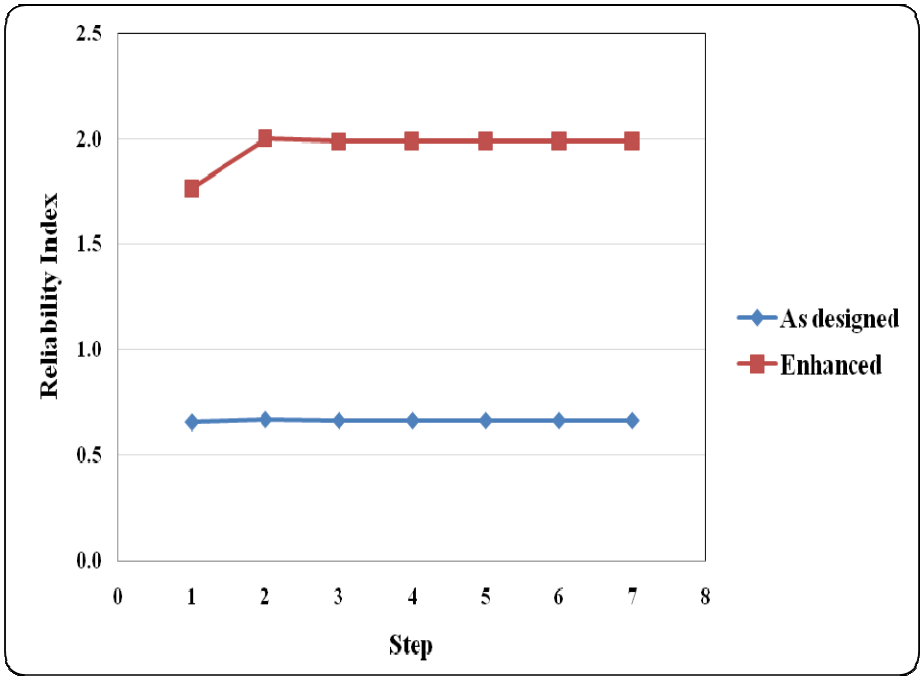


Figure 8.16 Reliability index in stepwise RSM iteration (Case study 2)

Table 8.12 Sensitivity analysis (Case study 2)

RANDOM VARIABLE	$\beta,_{U^*}$	$\beta,_{mean}$	$\beta,_{sd}$
E	2.504E-01	-4.017E-05	-5.462E-06
T_1	-1.311E-02	6.553E-02	-5.675E-04
T_2	8.108E-03	-2.134E-02	-1.143E-04
T_3	-8.997E-04	3.599E-03	-2.140E-06
T_4	2.339E-02	-6.496E-02	-1.004E-03
T_5	-1.023E-03	4.649E-03	-3.142E-06
T_6	-2.727E-04	8.523E-04	-1.536E-07
T_7	-1.184E-01	3.199E-01	-2.502E-02
T_8	1.438E-02	-4.108E-02	-3.904E-04
T_9	3.499E-04	-1.590E-03	-3.677E-07
T_{10}	-2.106E-06	6.582E-06	-9.162E-12
T_{11}	7.627E-02	-2.119E-01	-1.068E-02
K	-9.573E-01	8.128E-13	-5.532E-13

Table 8.13 Response surface (Case study 2)

RANDOM VARIABLE	**Coefficient**	**Partial Correlation** R_i
E	-3.424E-04	-7.797E-01
T_1	1.642E-01	2.882E-01
T_2	-2.574E-01	-6.679E-01
T_3	-1.720E-02	-1.158E-01
T_4	-1.357E-01	-4.094E-01
T_5	-1.507E-02	-8.952E-02
T_6	-1.518E-02	-1.306E-01
T_7	-3.862E+00	-2.549E-01
T_8	-1.632E-01	-4.643E-01
T_9	-2.855E-02	-1.679E-01
T_{10}	-1.705E-02	-1.464E-01
T_{11}	-1.998E+00	-8.996E-02
K	7.676E-12	9.315E-01
E^2	7.027E-10	9.999E-01
T_1^2	-5.775E-03	-5.308E-01
T_2^2	6.486E-03	9.304E-01
T_3^2	9.476E-04	1.585E-01
T_4^2	2.085E-03	5.910E-01
T_5^2	1.066E-03	1.385E-01
T_6^2	5.223E-04	1.435E-01
T_7^2	2.913E-02	9.958E-01
T_8^2	3.625E-03	7.693E-01
T_9^2	1.167E-03	1.514E-01
T_{10}^2	5.333E-04	1.465E-01
T_{11}^2	2.591E-02	9.940E-01
K^2	8.933E-27	9.979E-01
T_7E	8.540E-06	4.942E-01
$T_{11}E$	-4.537E-06	-1.911E-01
$T_{11}T_7$	8.804E-02	9.219E-02
KE	-2.809E-17	-9.896E-01
KT_1	1.986E-14	1.344E-01
KT_2	-7.881E-15	-1.017E-01
KT_4	-1.615E-14	-1.947E-01
KT_8	-1.079E-14	-1.279E-01
Regression Constant = 0.8521E2		

Table 8.14 Comparison with SORM and MCS (As designed)

	FORM	**SORM**	**MCS**
β	0.6609	0.6633	0.6631
P_f	0.2543	0.2536	0.2536

SORM: Direct integration based on curvature fitting
MCS: 10^8 samples drawn at design point, cov = 1.11E-4

Table 8.15 Comparison with SORM and MCS (Enhanced)

	FORM	**SORM**	**MCS**
β	1.9887	1.9932	1.9930
P_f	0.02337	0.02312	0.02313

SORM: Direct integration based on curvature fitting
MCS: 10^8 samples drawn at design point, cov = 1.53E-4

8.8 Conclusions

- Pseudo excitation method originally proposed by Lin (1992) in random vibration is employed to tackle non-linearity due to external wave pressure and inertia loading in fatigue assessment of ship structures. It is equally suitable for direct load method and unit load method in existing spectral fatigue frameworks.

- Though the discussion is based on the quasi-static system the origination of the approach hints that in a linear equivalent system hot spot stress can even be obtained through a proper dynamic analysis, which may well pick up in the future because of increasing power and plunging cost of computing technology.

- There are encouraging signs that step-wise RSM can be integrated into traditional spectral fatigue assessment process to describe fatigue damage in reliability terms.

- The boundary displacements at the mean design values are applied to the sub-model during fitting of response surface. Strictly speaking any change to the local scantlings will cause change in the boundary displacements and redistribution of the boundary reaction forces. To reduce their effect the sub-model includes a much bigger extent than the hotspot area. Moreover in each sample point only the thickness of one plate strake is changed by a small amount. So local loads and redistribution of local stiffness play the primary role in the stress change in the hot spot area. As another precaution, after the design is improved according to sensitivity information in the bulk carrier case study, global FE model is rerun to update the boundary displacements in the follow-up reliability calculation.

- Under the appearance of a highly complex spectral fatigue analysis there might be a fairly simple failure surface that can be approximated reasonably even with a linear RSF.

- The fatigue behaviour varies from location to location. So it is not sensible to project the conclusions drawn from the two case studies beyond the limits and assumptions in the current models.

- Only the effect of local scantlings is reflected in the model whereas in reality the increase of the section modulus can significantly reduce the hotspot stress at inner bottom to stool connection.

- Operational factors such as loading condition and trading pattern can severely affect fatigue life as well.

- The bulk carrier in case 2 was designed before Common Structural Rules (CSR 2004) came into force. CSR has more stringent fatigue criteria. In many cases the designers have gone as far as adding partial girders in the double bottom and applying weld treatment to meet the requirements. It can be expected that all this can give much higher fatigue reliability.

- The focus in modern ship design is increasingly on best weight distribution of the hull and optimised detail design with rule requirements and building cost in mind. The goals in different disciplines are often conflicting. The sensitivity information obtained through RSM at given structural detail can help designers make more balanced decision. This benefit comes before the still fledging reliability criteria become a commonplace in rules and regulations one day.

- Computing tools play an important role in evaluating reliability for large and complicated structural problems. Though they are not a decisive factor, it is foreseen that as more software systems in this area are available acceptance of reliability approach will certainly be accelerated. In addition algorithms with parallelization will surge and play a vital role in fatigue design.

References

1. Brard, R. (1972). The Representation of a given ship form by singularity distribution when the boundary on the free surface is linerised. *Journal of ship research*, 16(1).
2. British Marine Technology. (1986). *Global wave statistics*. BMT.

3. Chang, M. (1977). Computations for three dimensional ship motions with forward speed. *Second International Conference on Numerical Hydrodynamics*, University of California, Berkeley.

4. Dirlik, T. (1985). *Application of computers to fatigue analysis*. PhD Thesis, Warwick University. Warwick.

5. Det Norske Veritas. (2003). *Fatigue assessment of ship structures* (Classification Notes No. 30.7). DNV.

6. Folsø, R. (1998). Spectral fatigue damage calculation in the side shells of ships, with due account taken of the effect of alternating wet and dry areas. *Marine Structures* 11(7-8), 319-343.

7. Hansen, P.F. (1994). *Reliability analysis of a midship section*. Phd Thesis, Department of Naval Architecture & Offshore Engineering, Technical University of Denmark.

8. Korvin-Kroukovski, B. V., & Jacob, W. R. (1957). Pitching and heavy motions of ship in regular waves. *Transaction SNAME*. 65, 1957.

9. Lin, J.H. (1992). A fast CQC algorithm of PSD matrices for random seismic responses. *Computers & Structures.* 44(3), 683-687.

10. Lloyd's Register. (2004). *Fatigue Design Assessment Level 3 Procedure*. LR.

11. Lloyd's Register. (2002). *ShipRight Fatigue Design Assessment Level 2: User Manual*.

12. MARIN. (2003). *Precal 6.0 Theory Manual*.

13. Newman, J. N. (1970). Application of slender body theory in ship hydrodynamics. *Annual Review of Fluid Mechanics*, 2, 67-94.

14. Petinov, S. (2003). *Fatigue analysis of ship structures*. New Jersey: Backbone Publishing Company.

15. Sherratt, F., Bishop, N.W.M., & Dirlik, T. (2005). Predicting fatigue life from frequency domain data. *Engineering Integrity* 18, 12-16.

16. Soares, C.G., & Moan, T. (1991). Model uncertainty in the long term distribution of wave induced bending moments for fatigue design of ship structures. *Marine Structures*, 4(4), 295-315.

17. Violette, F.L.M. (1997). *On the fatigue performance prediction of ship structural details*. PhD Thesis, University of Southampton, Southampton.

18. Wehausen, J. V., & Llitone, B. V. (1960). Surface waves. *Handbuch der Physik* (pp. 446-778), Berlin: Springer-Verlag.

19. Winterstein, S. R. (1988). Nonlinear vibration models for extremes and fatigue. *Journal of Engineering Mechanics*, ASCE, 114(10), 1772-1790.

20. Wirsching, P.H. & Shehata, A.M. (1977). Fatigue under wide-band random stresses using rainflow method. *Journal of Engineering Materials and Technology*, 99(3), 205-211.

21. Zhou. M, et al. (2009). Spectral Fatigue Analysis of a Bulk Carrier – A Case Study. *Proceedings of Design and Operation of Bulk Carriers*, RINA.

Chapter 9 Conclusions and Prospects

9.1 Achievements

A fatigue reliability framework has been established and tested in the current research. It contains a coherent core of second moment reliability methods, simulation techniques, and response surface method. On top of that is the spectral fatigue model based on the first principle sea-keeping and structural analysis. More complicated problems such as system reliability and reliability based optimization can be broken down into nested component reliability problems. Meanwhile more advanced fatigue models, such as those based on fracture mechanics, need to be fully developed to be of practical use. So the current framework allows much room for further expansion. The main achievements are summarized below:

- Non-Gaussian random process generation through zero mean non-linear transformation of Gaussian process is studied with FFT acceleration. Hermite expansion is used to convert the autocorrelation function and obtain the corresponding spectrum for the underlying Gaussian process. Impact of time interval on the generated process is discussed in numerical examples.

- A new set of formulations for second moment reliability theory is derived. Compared to existing formulations in the text books the new ones are more generic. Details on how to treat correlated random variables with arbitrary distribution are revealed step by step. It is proven that traditional FORM methods such as RF-HL are special cases of the quadratic programming problem. Different asymptotic integration and curvature fitting techniques are discussed in an extensive comparison between SORM methods. The precise integration of quadratic form in the normal space is used as the baseline to compare with. It can be evaluated by Fourier transform of the characteristic function.

- Monte Carlo simulation techniques are reviewed. A new Markov chain kernel method is proposed as a variation of importance sampling method. It requires no pre-knowledge of the limit state function itself. Optimal kernel density can be built up adaptively to approximate the unknown theoretical sampling density in the failure region. The method is applied to a FEM double bottom model to verify the results from response surface method.

- Step-wise response surface method is proposed. It allows terms in the quadratic response function to be selected according to their statistical contributions. The process can avoid singularity problems as in the standard regression since if the role of a term is already represented by selected terms it will not be picked up. Sampling techniques are discussed to maximise reusability in the adaptive iteration. Meanwhile any term can be forced to be included or excluded through user defined weighting factors from engineering judgement.

- A systematic approach to study the effect of both non-Gaussianness and bandwidth of the stress response is proposed. This is an extension of the well known Wirsching model, which can only deal with bandwidth effect. A range of skewness and kurtosis is studied by employing distributions in the Johnson family. The correction curves are presented in a format suitable for rules and procedures.

- Literature review of time variant problem is given. Ways to convert it to nested reliability problems are summarized. This provides a growth point for the framework in the future research.

- Pseudo excitation method is introduced to traditional spectral fatigue model. It provides better theoretical grounds than a simple proscriptive wave height used to get linear equivalent stress RAO. Stepwise response surface method is further applied to this model successfully in a tanker and bulk carrier case study. Not only can this combination be used to calibrate the reliability level of existing designs but it also gives informative sensitivity results that lead to improved design.

9.2 Conclusions

Conclusions at algorithm level are already drawn in individual chapters. So a helicopter view is offered here to generalize lessons learned for future research.

The adoption of modelling, transformation and optimization as the basic methodologies for this research has proven its worth. Modelling is the starting point. Transformation provides specific view angles that can simplify the problem in hand. And optimization is the means for best point of solution. For instance, when deriving second moment formulations the original problem is first transformed to correlated normal space and then uncorrelated standard normal space, where optimization methods are used to determine the reliability index, namely the minimum distance. Construction of kernel density in Monte Carlo simulation involves modelling the kernel and optimizing it according to predefined criteria. Same is true in response surface method.

The devil is in the details. By fully digesting developed numerical methods one can often see ways to improve emerging methods. Stepwise response surface method was proposed in this way when the author was studying stepwise regression. Had the author used solution from existing mathematical and statistical library he would not have come up such an idea at all. Other proposed methods in the research are natural result of reproducing existing methods in computer program first and then finding refinements to solve the problems encountered.

Structural reliability analysis (RSA) belongs to Bayesian probability realm. Different viewpoint will give different reliability estimate of the same problem. Engineering judgment is essential to distinguish primary factors from secondary ones during modelling. A good reliability model is often not the most complete one but the one with better usability and can be easily calibrated and updated using data collected from experience.

Though reliability methods have been used in background calibration of the rules and procedures the awareness of which among ship designers is still pretty low. Reliability practitioners face a big challenge to help key stake holders to understand the benefit of

using RSA towards safer yet more balanced design. As is shown in the thesis sensitivity analysis can provide valuable design information to the designer's best appeal.

Lack of common codes in the industry is still the major hindrance of application of reliability methods. Many codes are in the form of guidance, which is not compulsory. In the foreseeable future traditional prescriptive standards in ship building industry will gradually give way to more transparent goal-based standards (GBS) to encourage innovations that can better mitigate risks to life at sea and environment. Since risk-based principles and the underlying probability language can be applied across the components and subsystems of the overall ship and human system, they may change the skyline of rules and procedures in the decades to come, just as direct calculation once did.

It is the author's belief that for reliability criteria to find a way in the rules and procedures the whole industry must try a lot in collaboration and software will no doubt speed up the iteration process.

9.3 GLAREL

The reliability methods covered in this thesis have been implemented by the author in GLAREL software, which is used to provide results for all the examples and case studies at the end of each chapter. The software has also been used as a key tool by many fellow PhD students at the University of Glasgow and Strathclyde in their research since 2001.

GLAREL is written in FORTRAN language. It fully supports FORM, SORM, MCS and RSM. Data input is through text file with clearly separated data blocks bounded by the keywords. The main features are listed in Table 9.1:

Table 9.1 Features available in GALREL

Category		**Options**
FORM		HL-RF
		Advanced HL-RF
		BFGS SQP
		Newton SQP (Hessian matrix required)
		NLSQP
SORM	Fitting	Curvature fitting
		Point fitting
	Quadratic form	Parabola
		Taylor expansion
	Integration	Asymptotic
		Direct integration
MCS		Multi-normal adaptive sampling
		Kernel adaptive sampling
RSM		Stepwise RSM

9.4 Prospects

The current study makes the author better positioned to take on future researches in the following fields:

- Applying time-variant fatigue model to longitudinal end connection problems based on beam theory, stress concentration factor library and parametric hydro loads. This will be followed by time-variant model for fatigue damage based on the first principles.

- Considering fatigue damage together with yielding and buckling failures in a system reliability context. These are three most important areas to check in ship structural design and construction. Though individually reliability methods may have been used to calibrate the criteria in the rules and procedures their totality effect based on system reliability has yet to be revealed for critical local designs.

- Studying risk-based inspection models based on reliability results and their application in hull survey and maintenance.

- Exploring reliability based optimization techniques and their application in structural design within the boundary of rules and regulations.

- Enhancing software development to fully leverage the available computing power and bring down the cost of reliability methods. Experience shows this is the most effective way to facilitate knowledge transfer and boost productivity.

9.5 Publications

The following papers have been published as a result of the current research:

1. Yu, L., Das, P. K., & Barltrop, N. D. (2001). Importance sampling method with kernel density estimate – A general approach. *Proceedings of the 8th International Conference on Structural Safety and Reliability* (ICOSSAR 2001), California.

2. Yu, L., Das, P. K., & Zheng, Y. L. (2002). Stepwise response surface method and its application to reliability analysis of ship structures. *Journal of Offshore Mechanics and Arctic Engineering*, 124(4), 226-230.

3. Yu, L., Das, P. K., & Barltrop, N. D. (2004). A new look at the effect of bandwidth and non-normality on fatigue damage. *Fatigue & Fracture of Engineering Material & Structures*, 27(1), 51-58.

4. Yu, L., & Das, P. K. (2007). Fatigue design assessment based on pseudo-excitation method. *International Maritime Association of Mediterranean* (IMAM), Varna.

5. Yu, L., Das, P. K., & Zheng, Y. L. (2009). A response surface approach to fatigue reliability of ship structures. *Ship and Offshore Structures*, 4(3), 253-259.

6. Yu. L., Zhou M., Das P. K., et al. (2010). Fatigue reliability of bulk carrier – A case study. *Advanced Structural Reliability Analysis Network*, Edinburgh.